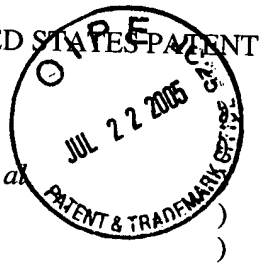
B/ERD

PATENT
Attorney Docket No. 62785.000005

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE



In re Application of:

LE PAGE, RICHARD, *et al.*

Serial Number: 09/769,736

Filed: January 26, 2001

For: NUCLEIC ACIDS AND PROTEINS FROM GROUP B STREPTOCOCCUS

Group Art Unit: 1637

Examiner: K. CARLSON

SUBMISSION OF COPY OF FOREIGN PRIORITY DOCUMENT UNDER 35 U.S.C. 119(a)-(d)

Mail Stop Amendment
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Examiner Carlson:

Further to our telephone conversation of July 5, 2005, submitted herein is a copy of the Foreign Priority document for the instant Application (GB 9816335.5).

No fees are believed to be due in connection with the entry of the instant document. The Commissioner is hereby authorized to charge any fees which may be required, or credit any overpayment, to Deposit Account Number 50-0206.

Respectfully submitted,
HUNTON & WILLIAMS LLP

Dated: 7/22/05

By: _____
Laurence H. Posorske, Ph.D.
Registration No. 34,698

Christopher J. Nichols, Ph.D.
Registration No. 55,984

Hunton & Williams LLP
Intellectual Property Department
1900 K Street, N.W., Suite 1200
Washington, DC 20006-1109
(202) 955-1500 (telephone)
(202) 778-2201 (facsimile)
LHP/CJN:cdh

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

I also certify that the attached copy of the request for grant of a Patent (Form 1/77) bears a correction, effected by this office, following a request by the applicant and agreed to by the Comptroller-General.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

CERTIFIED COPY OF PRIORITY DOCUMENT

Signed Stephen Hordley

Dated 2 July 2002

Patents Form 1/77

The Patent Office

1/77

28JUL98 E378889-1 D00056
P01/7700 25.00 - 9816335.5

Patents Act 1977
(Rule 16)

THE PATENT OFFICE
27 JUL 1998
NEWPORT

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

The Patent Office

Cardiff Road
Newport
Gwent NP9 1RH

1. Your reference

PWC/P21089GB

2. Patent application number
(The Patent Office will fill in this part)

27 JUL 1998

9816335.5

3. Full name, address and postcode of the or of each patent applicant *(underline all surnames)*

~~CORTECS UK LIMITED~~
~~THE OLD BLUE SCHOOL~~
~~LOWER SQUARE~~
~~ISLEWORTH~~
~~MIDDLESEX TW7 6RL~~
~~UK~~

Microbial Technics Limited
20, Trumpington Street
Cambridge
CB2 1QA.

CF 20/8/99.

~~2326023001~~
06832471001.

Patents ADP number *(if you know it)*

If the applicant is a corporate body, give the country/state of its incorporation

4. Title of the invention

PROTEINS

5. Name of your agent *(if you have one)*

"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*

KILBURN & STRODE
20 RED LION STREET
LONDON
WC1R 4PJ

Patents ADP number *(if you know it)*

125001

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or each of these earlier applications and *(if you know it)* the or each application number

Country	Priority application number *(if you know it)*	Date of filing *(day / month / year)*

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application number	Date of filing *(day / month / year)*

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer 'Yes' if:*

YES

a) any applicant named in part 3 is not an inventor, or

b) there is an inventor who is not named as an applicant, or

c) any named applicant is a corporate body. See note (d))

Patents Form 1/77

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form	:	
Description	:	21
Claim(s)	:	4
Abstract	:	
Drawing(s)	:	29

10. If you are also filing any of the following, state how many against each item.

Priority documents	:
Translations of priority documents	:
Statement of inventorship and right to grant of a patent *(Patents Form 7/77)*	:
Request for preliminary examination and search *(Patents Form 9/77)*	:
Request for substantive examination *(Patents form 10/77)*	:
Any other documents *(please specify)*	:

11. I/We request the grant of a patent on the basis of this application.

Signature: Kilburn & Strode

Date: 27 July 1998

12. Name and daytime telephone number of person to contact in the United Kingdom

Paul Chapman
Tel: 0171-242 8291

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505.

b) Write your answers in capital letters using black ink or you may type them.

c) If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.

d) If you have answered `Yes' Patents Form 7/77 will need to be filed.

e) Once you have filled in the form you must remember to sign and date it.

f) For details of the fee and ways to pay please contact the Patent Office.

Proteins

The present invention relates to proteins derived from *Streptococcus agalactiae*, nucleic acid molecules encoding such proteins, and the use of the proteins as antigens and/or immunogens and in detection/diagnosis. It also relates to a method for the rapid screening of bacterial genomes to isolate and characterise bacterial cell envelope associated or secreted proteins.

The *Group B Streptococcus* (GBS) *(Streptococcus agalactiae)* is an encapsulated bacterium which emerged in the1970s as a major pathogen of humans causing sepsis and meningitis in neonates as well as adults. The incidence of early onset neonatal infection during the first 5 days of life varies from 0.7 to 3.7 per 1000 live births and causes mortality in about 20% of cases. Between 25-50% of neonates surviving early onset infections frequently suffer neurological sequalae. Late onset neonatal infections occur from 6 days to three months of age at a rate of about 0.5 - 1.0 per 1000 live births.

There is an established association between the colonisation of the maternal genetic tract by GBS at the time of birth and the risk of neonatal sepsis. In humans it has been established that the rectum may act as a reservoir for GBS. Susceptibility in the neonate is correlated with the a low concentration or absence of IgG antibodies to the capsular polysaccharides found on GBS causing human disease. In the USA strains isolated from clinical cases usually belong to capsular serotypes Ia, Ib, II, III although serotype V may be of increasing significance. Type VIII GBS is the major cause of neonatal sepsis in Japan.

A possible means of prevention involves intra or postpartum administration of antibiotics to the mother but there are concerns that this might lead to the emergence of resistant organisms and in some cases allergic reactions. Vaccination of the adolescent females to induce long lasting maternally derived immunity is one of the most promising approaches to prevent GBS infections in neonates. The capsular

polysaccharide antigens of these organisms have attracted most attention as with regard to vaccine development. Studies in healthy adult volunteers have shown that serotype 1a, II and III polysaccharides are non-toxic and immunogenic in approximately 65%, 95% and 70% of non-immune adults respectively. One of the problems with using capsule antigens as vaccines is that the response rates vary according to pre-immunisation status and the polysaccharide antigen and not all vaccinees produce adequate levels of IgG antibody as indicated in vaccination studies with GBS polysaccharides in human volunteers.

Some people do not respond despite repeated stimuli. These properties are due to the T-independent nature of polysaccharide antigens. One strategy to enhance the immunogenicity of these vaccines is to enhance the T cell dependent properties of polysaccharides by conjugating them to a protein. The use of polysaccharide conjugates looks promising but there are still unresolved questions concerning the nature of the carrier protein. A conjugate vaccine against GBS would require at least 4 different conjugates to be prepared adding to the cost of a vaccine.

Recent evidence also suggests that bacterial surface proteins may be useful to confer immunity. A protein called Rib which is found on most serotype III strains but rarely on serotypes Ia, Ib or II confers immunity to challenge with Rib expressing GBS in animal models (Stalhammar-Carlemalm *et al., Journal of Experimental Medicine* 177:1593-1603 (1993)). Another surface protein of interest as a component of a vaccine is the alpha antigen of the C proteins which protected vaccinated mice against lethal infection with strains expressing alpha protein. The amount of antigen expressed by GBS strains varies markedly.

Approaches to vaccination against GBS infections which rely on the use of capsular polysaccharides have the disadvantage that response rates are likely to vary considerably according to pre-immunisation status and the particular type of polysaccharide antigen used. Results of trials in human volunteers have indicated that

response rates may only be around 65% for some of the key capsule antigens (Larsson *et al.*, *Infection and Immunity* **64**:3518-3523 (1996)). It is also not clear whether all individuals responding to the vaccine would have adequate levels of polysaccharide specific IgG which can cross the placenta and afford immunity to neonates. By conjugating a protein carrier to the polysaccharide antigen it may be possible to convert them to T-cell dependent antigens and enhance their immunogenicity.

Preliminary studies with GBS type III polysaccharide-tetanus toxoid conjugate have been encouraging (Baker *et al.*, *Reviews of Infectious Diseases* **7**:458-467 (1985), Baker *et al.*, *The New England Journal of Medicine* **319**:1180-1185 (1988), Paoletti *et al.*, *Infection and Immunity* **64**:677-679 (1996), Paoletti *et al.*, *Infection and Immunity* **62**:3236-3243 (1994)) but in developed countries the use of tetanus may be disadvantageous since most adults will have been immunised against tetanus within the past five years. Additional boosters with tetanus toxoid may cause adverse reactions (Boyer., *Current Opinions in Pediatrics* **7**:13-18 (1995)). The polysaccharide conjugate vaccines have the disadvantage of being costly to produce and manufacture in comparison with many other kinds of vaccines. There is also the possible risk of problems caused by the cross reactivity between GBS polysaccharides and sialic acid-containing human glycoproteins.

An alternative to polysaccharides as antigens is the use of protein antigens derived from GBS. Recent evidence suggest that the GBS surface associated proteins Rib and alpha C protein may be used to confer immunity to GBS infections in experimental model systems (Stalhammar-Carlemalm *et al.*, (1993) [*supra*], Larsson *et al.*, (1996) [*supra*]). However these two proteins are not conserved in all serotypes of GBS which cause disease in humans. Assuming that these antigens would be immunogenic and elicit protective level responses in humans they would not confer protection against all infections as 10% of infectious *Group B streptococci* do not express Rib or C protein alpha.

This invention seeks to overcome the problem of vaccination against GBS by using a novel screening method specifically designed to identify those *Group B Streptococcus* genes encoding bacterial cell surface associated or secreted proteins (antigens). The proteins expressed by these genes may be immunogenic, and therefore may be useful in the prevention and treatment of *Group B Streptococcus* infection. For the purposes of this application, the term immunogenic means that these proteins will elicit a protective immune response within a subject. Using this novel screening method a number of genes encoding novel *Group B Streptococcus* proteins have been identified.

Thus in a first aspect, the present invention provides a *Group B Streptococcus* protein, having a sequence selected from those shown in figure 1, or fragments or derivatives thereof.

In a further aspect, the present invention provides a *Group B Streptococcus* polypeptide or peptide having a sequence selected from those shown in figure 2, or fragments or derivatives thereof.

It will be apparent to the skilled person that proteins and polypeptides included within this group may be cell surface receptors, adhesion molecules, transport proteins, membrane structural proteins, and/or signalling molecules.

Alterations in the amino acid sequence of a protein can occur which do not affect the function of a protein. These include amino acid deletions, insertions and substitutions and can result from alternative splicing and/or the presence of multiple translation start sites and stop sites. Polymorphisms may arise as a result of the infidelity of the translation process. Thus changes in amino acid sequence may be tolerated which do not affect the proteins function.

Thus, the present invention includes derivatives or variants of the proteins, polypeptides, and peptides of the present invention which show at least 50% identity

to the proteins, polypeptides and peptides described herein. Preferably the degree of sequence identity is at least 60% and preferably it is above 75%. More preferably still is it above 80%, 90% or even 95%.

The term identity can be used to describe the similarity between two polypeptide sequences. A software package well known in the art for carrying out this procedure is the CLUSTAL program. It compares the amino acid sequences of two polypeptides and finds the optimal alignment by inserting spaces in either sequence as appropriate. The amino acid identity or similarity (identity plus conservation of amino acid type) for an optimal alignment can also be calculated using a software package such as BLASTx. This program aligns the largest stretch of similar sequence and assigns a value to the fit. For any one pattern comparison several regions of similarity may be found, each having a different score. One skilled in the art will appreciate that two polypeptides of different lengths may be compared over the entire length of the longer fragment. Alternatively small regions may be compared. Normally sequences of the same length are compared for a useful comparison to be made.

Manipulation of the DNA encoding the protein is a particularly powerful technique for both modifying proteins and for generating large quantities of protein for purification purposes. This may involve the use of PCR techniques to amplify a desired nucleic acid sequence. Thus the sequence data provided herein can be used to design primers for use in PCR so that a desired sequence can be targeted and then amplified to a high degree.

Typically primers will be at least five nucleotides long and will generally be at least ten nucleotides long (e.g. fifteen to twenty-five nucleotides long). In some cases primers of at least thirty or at least thirty-five nucleotides in length may be used.

As a further alternative chemical synthesis may be used. This may be automated. Relatively short sequences may be chemically synthesised and ligated together to provide a longer sequence.

Thus in a further aspect, the present invention provides , a nucleic acid molecule comprising or consisting of a sequence which is:

(i) any of the DNA sequences set out in figure 1 or figure 2 herein or their RNA equivalents;

(ii) a sequence which is complementary to any of the sequences of (i);

(iii) a sequence which codes for the same protein or polypeptide, as those sequences of (i) or (ii);

(iv) a sequence which is shows substantial identity with any of those of (i), (ii) and (iii); or

(v) a sequence which codes for a derivative or fragment of a nucleic acid molecule shown in figure 1 or figure 2.

The term identity can also be used to describe the similarity between two individual DNA sequences. The 'bestfit' program (Smith and Waterman, Advances in applied Mathematics, 482-489 (1981)) is one example of a type of computer software used to find the best segment of similarity between two nucleic acid sequences, whilst the GAP program enables sequences to be aligned along their whole length and finds the optimal alignment by inserting spaces in either sequence as appropriate.

The term 'RNA equivalent' when used above indicates that a given RNA molecule has a sequence which is complementary to that of a given DNA molecule, allowing for the fact that in RNA 'U' replaces 'T' in the genetic code. The nucleic acid molecule may be in isolated or recombinant form.

The nucleic acid molecule may be in an isolated or recombinant form. DNA constructs can readily be generated using methods well known in the art. These techniques are disclosed, for example in J. Sambrook *et al*, *Molecular Cloning 2nd Edition*, Cold Spring Harbour Laboratory Press (1989). Modifications of DNA constructs and the proteins expressed such as the addition of promoters, enhancers, signal sequences,

leader sequences, translation start and stop signals and DNA stability controlling regions, or the addition of fusion partners may then be facilitated.

Normally the DNA construct will be inserted into a vector which may be of phage or plasmid origin. Expression of the protein is achieved by the transformation or transfection of the vector into a host cell which may be of eukaryotic or prokaryotic origin. Such vectors and suitable host cells form yet further aspects of the present invention.

The *Group B Streptococcus* proteins (antigens) described herein can additionally be used to raise antibodies, or to generate affibodies. These can be used to detect *Group B Streptococcus.*

Thus in a further aspect the present invention provides, an antibody, affibody, or a derivative thereof which binds to any one or more of the proteins, polypeptides, peptides, fragments or derivatives thereof, as described herein.

Antibodies within the scope of the present invention may be monoclonal or polyclonal. Polyclonal antibodies can be raised by stimulating their production in a suitable animal host (e.g. a mouse, rat, guinea pig, rabbit, sheep, goat or monkey) when a protein as described herein, or a homologue, derivative or fragment thereof, is injected into the animal. If desired, an adjuvant may be administered together with the protein. Well-known adjuvants include Freund's adjuvant (complete and incomplete) and aluminium hydroxide. The antibodies can then be purified by virtue of their binding to a protein as described herein.

Monoclonal antibodies can be produced from hybridomas. These can be formed by fusing myeloma cells and spleen cells which produce the desired antibody in order to form an immortal cell line. Thus the well-known Kohler & Milstein technique (*Nature* **256** (1975)) or subsequent variations upon this technique can be used.

Techniques for producing monoclonal and polyclonal antibodies that bind to a particular polypeptide/protein are now well developed in the art. They are discussed in standard immunology textbooks, for example in Roitt *et al*, *Immunology* second edition (1989), Churchill Livingstone, London.

In addition to whole antibodies, the present invention includes derivatives thereof which are capable of binding to proteins etc as described herein. Thus the present invention includes antibody fragments and synthetic constructs. Examples of antibody fragments and synthetic constructs are given by Dougall *et al* in *Tibtech* **12** 372-379 (September 1994).

Antibody fragments include, for example, Fab, $F(ab')_2$ and Fv fragments. Fab fragments (These are discussed in Roitt *et al* [*supra*]). Fv fragments can be modified to produce a synthetic construct known as a single chain Fv (scFv) molecule. This includes a peptide linker covalently joining V_h and V_l regions, which contributes to the stability of the molecule. Other synthetic constructs that can be used include CDR peptides. These are synthetic peptides comprising antigen-binding determinants. Peptide mimetics may also be used. These molecules are usually conformationally restricted organic rings that mimic the structure of a CDR loop and that include antigen-interactive side chains.

Synthetic constructs include chimaeric molecules. Thus, for example, humanised (or primatised) antibodies or derivatives thereof are within the scope of the present invention. An example of a humanised antibody is an antibody having human framework regions, but rodent hypervariable regions. Ways of producing chimaeric antibodies are discussed for example by Morrison *et al* in *PNAS*, **81,** 6851-6855 (1984) and by Takeda *et al* in *Nature*. **314,** 452-454 (1985).

Synthetic constructs also include molecules comprising an additional moiety that provides the molecule with some desirable property in addition to antigen binding. For

example the moiety may be a label (e.g. a fluorescent or radioactive label). Alternatively, it may be a pharmaceutically active agent.

Affibodies are proteins which are found to bind to target proteins with a low dissociation constant. They are selected from phage display libraries expressing a segment of the target protein of interest (Nord K, Gunneriusson E, Ringdahl J, Stahl S, Uhlen M, Nygren PA, Department of Biochemistry and Biotechology, Royal Institute of Technology (KTH), Stockholm, Sweden).

In a further aspect the invention provides an immunogenic composition comprising one or more proteins, polypeptides, peptides, fragments or derivatives thereof, or nucleotide sequences described herein. A composition of this sort may be useful in the treatment or prevention of *Group B Streptococcus* infection in subject. In a preferred aspect of the invention the immunogenic composition is a vaccine.

In other aspects the invention provides:

i) Use of an immunogenic composition as described herein in the preparation of a medicament for the treatment or prophylaxis of *Group B Streptococcus* infection. Preferably the medicament is a vaccine.

ii) A method of detection of *Group B Streptococcus* which comprises the step of bringing into contact a sample to be tested with at least one antibody, affibody, or a derivative thereof, as described herein.

iii) A method of detection of *Group B Streptococcus* which comprises the step of bringing into contact a sample to be tested with at least one protein, polypeptide, peptide, fragments or derivatives as described herein.

iv) A method of detection of *Group B Streptococcus* which comprises the step of bringing into contact a sample to be tested with at least one nucleic acid molecule as described herein.

v) A kit for the detection of *Group B Streptococcus* comprising at least one antibody, affibody, or derivatives thereof, described herein.

vi) A kit for the detection of *Group B Streptococcus* comprising at least one *Group B Streptococcus* protein, polypeptide, peptide, fragment or derivative thereof, as described herein.

vii) A kit for the detection of *Group B Streptococcus* comprising at least one nucleic acid of the invention.

As described previously, the novel proteins described herein are identified and isolated using a novel screening method which specifically identifies those *Group B Streptococcus* genes encoding bacterial cell envelope associated or secreted proteins.

The information necessary for the secretion/export of proteins has been extensively studied in bacteria. In the majority of cases, export requires a signal peptide positioned at the N-terminus of the precursor protein to target the precursor to translocation sites on the membrane. During or after translocation, the signal peptide is removed by a signal peptidase. The ultimate destination/localisation of the protein, (whether it be secreted extracellularly or anchored to the bacterium's surface, etc) is determined by sequences other than the leader peptide sequence.

Recently, Poquet *et al.* (*J. Bacteriol.* **180:**1904-1912 (1998)) have described a screening vector incorporating the *nuc* gene lacking its own signal leader as a reporter to identify exported proteins in Gram positive bacteria, and have applied it to *L. lactis*.

Staphylococcal nuclease is a naturally secreted heat-stable, monomeric enzyme which has been efficiently expressed and secreted in a range of Gram positive bacteria (Shortle., *Gene* **22:**181-189 (1983), Kovacevic *et al.*, *J. Bacteriol.* **162:**521-528 (1985), Miller *et al.*, *J. Bacteriol.* **169:**3508-3514 (1987), Liebl *et al.*, *J. Bacteriol.* **174:**1854-1861(1992), Le Loir *et al.*, *J. Bacteriol.* **176:**5135-5139 (1994), Poquet *et al.*, 1998 [*supra*]). The screening vector (pFUN) contains the pAMβ1 replicon which functions in a broad host range of Gram-positive bacteria in addition to the ColE1 replicon that promotes replication in *Escherichia coli* and certain other Gram negative bacteria. Unique cloning sites present in the vector can be used to generate transcriptional and translational fusions between cloned genomic DNA fragments and the open reading frame of the truncated *nuc* gene devoid of its own signal secretion leader. The *nuc* gene makes an ideal reporter gene because the secretion of nuclease can readily be detected using a simple and sensitive plate test: Recombinant colonies secreting the nuclease develop a pink halo whereas control colonies remain white (Shortle, 1983 [*supra*], Le Loir *et al.*, 1994 [*supra*]).

A direct screen to identify and isolate DNA encoding bacterial cell envelope associated or secreted proteins (antigens).in pathogenic bacteria has been developed by the present inventors which utilises a vector-system (pTREP1 expression vector) in *Lactococcus lactis* that specifically detects DNA sequences which are adjacent to, and associated with DNA encoding surface proteins from *Group B Streptococcus*. The screening vector also incorporates the *nuc* gene encoding the *Staphylococcal* nuclease as a reporter gene.

Only the part of the *nuc* gene encoding the mature nuclease protein (minus its signal peptide sequence) is cloned into the pTREP1 expression vector in *L. lactis*. In this form, the *nuc*-encoded nuclease cannot be secreted even when expressed intracellularly. The reporter vector is then randomly combined with appropriately digested genomic DNA from *Group B Streptococcus*, cloned into *L. lactis* and used as

a screening system for sequences permitting the export of nuclease. In this way gene/partial gene sequences encoding exported proteins from *Group B Streptococcus* are isolated. Once a partial gene sequence is obtained, full length sequences encoding exported proteins can readily be obtained using techniques well known in the art.

In possessing a promoter, the pTREP1-*nuc* vectors differ from the pFUN vector described by Poquet *et al.* (1998) [*supra*], which was used to identify *L. lactis* exported proteins by screening directly for *Nuc* activity directly in *L. lactis*. As the pFUN vector does not contain a promoter upstream of the *nuc* open reading frame the cloned genomic DNA fragment must also provide the signals for transcription in addition to those elements required for translation initiation and secretion of *Nuc*. This limitation may prevent the isolation of genes that are distant from a promoter for example genes which are within polycistronic operons. Additionally there can be no guarantee that promoters derived from other species of bacteria will be recognised and functional in *L. lactis*. Certain promoters may be under stringent regulation in the natural host but not in *L. lactis*. In contrast, the presence of the P1 promoter in the pTREP1-*nuc* series of vectors ensures that promoterless DNA fragments (or DNA fragments containing promoter sequences not active in *L. lactis*) may still be transcribed. Thus yet another advantage of this invention is that genes missed in other screening methods may be identified.

Hence in a further aspect the present invention provides a method of screening for DNA encoding bacterial cell wall associated or surface antigens in gram positive bacteria comprising the steps of:

- combining a reporter vector including the nucleotide sequence encoding the mature from of the staphylcoccus nuclease gene and an upstream promoter region with DNA from a gram positive bacteria.
- transforming the resultant vector into *Lactococcus lactis* cells.
- assaying for the secretion of *staphlycoccus* nuclease protein in the transformed cells.

Preferably, the reporter vector is one of the pTREP1-*nuc* vectors shown in figure 4.

In another aspect, the present invention provides a vector as shown in figure 4 for use in screening for DNA encoding exported or surface antigens in gram positive bacteria. Examples of gram positive bacteria which may be screened include *Group B Streptococcus*, *Streptococcus pneumoniae*, *Staphylcoccus aureus or pathogenic Group A Streptococci.*

Given that the inventors have identified a group of important proteins, such proteins are potential targets for anti-microbial therapy. It is necessary, however, to determine whether each individual protein is essential for the organism's viability. Thus, the present invention also provides a method of determining whether a protein or polypeptide as described herein represents a potential anti-microbial target which comprises inactivating said protein and determining whether *Group B Streptococcus* is still viable.

A suitable method for inactivating the protein is to effect selected gene knockouts, ie prevent expression of the protein and determine whether this results in a lethal change. Suitable methods for carrying out such gene knockouts are described in Li *et al*, *P.N.A.S.*, **94**:13251-13256 (1997) and Kolkman *et al*

In a final aspect the present invention provides the use of an agent capable of antagonising, inhibiting or otherwise interfering with the function or expression of a protein or polypeptide of the invention in the manufacture of a medicament for use in the treatment or prophylaxis of *Group B Streptococcus* infection.

The invention will now be described by means of the following example which should not in any way be construed as limiting. The examples refer to the figures in which

Fig 1: (A) Shows a number of full length nucleotide sequences encoding antigenic *Group B Streptococcus* proteins. (B) Shows the corresponding amino acid sequences.

Fig 2: (A) Shows a number of partial nucleotide sequences encoding antigenic *Group B Streptococcus* polypeptides and peptides. (B) Shows the corresponding amino acid sequences.

Fig 3: Shows a number of oligonucleotide primers used in the screening process

nucS1 primer designed to amplify a mature form of the nuc A gene

nucS2- primer designed to amplify a mature form of the nuc A gene.

nucS3 primer designed to amplify a mature form of the nuc A gene

nucR primer designed to amplify a mature form of the nuc A gene

nucseq primer designed to sequence DNA cloned into the pTREP-Nuc vector

pTREPF nucleic acid sequence containing recognition site for ECORV. Used for cloning fragments into pTREX7.

pTREPR nucleic acid sequence containing recognition site for BAMH1. Used for cloning fragments into pTREX7.

PUCF forward sequencing primer, enables direct sequencing of cloned DNA fragments.

VR example of gene specific primer used to obtain further antigen DNA sequence by the method of DNA walking.

V1 example of gene specific primer used to obtain further antigen DNA sequence by the method of DNA walking.

V2 example of gene specific primer used to obtain further antigen DNA sequence by the method of DNA walking.

Fig 4: (i) Schematic presentation of the nucleotide sequence of the unique gene cloning site immediately upstream of the mature *nuc* gene in pTREP1-*nuc*1, pTREP1-*nuc*2 and pTREP1-*nuc*3. Each of the pTREP-*nuc* vectors contain an EcoRV (a SmaI site in pTREP1-*nuc*2) cleavage site which allows cloning of genomic DNA fragments in 3 different frames with respect to the mature *nuc* gene.

(ii) A physical and genetic summary map of the pTREP1-*nuc* vectors. The expression cassette incorporating *nuc*, the macrolides, lincosamides and streptogramin B (MLS) resistance determinant, and the replicon (rep) *Ori*-pAMβ1 are depicted (not drawn to scale).

(iii) Schematic presentation of the expression cassette showing the various sequence elements involved in gene expression and location of unique restriction endonuclease sites (not drawn to scale).

Example 1

Thus far more than 100 gene/partial gene sequences putatively encoding exported proteins in *S. agalactiae* have been identified using the nuclease screening system of the invention. These have been further analysed to remove artifacts. The nucleotide sequences of genes identified using the screening system has been characterised using a number of parameters described below. All of these sequences are novel in that they have not been described previously.

1. All putative surface proteins are analysed for leader/signal peptide sequences. Bacterial signal peptide sequences share a common design. They are characterised by a short positively charged N-terminus (N region) immediately preceding a stretch of hydrophobic residues (central portion-h region) followed by a more polar C-terminal portion which contains the cleavage site (c-region). Computer software is used to perform hydropathy profiling of putative proteins (Marcks, *Nuc. Acid. Res.*, **16**:1829-1836 (1988)) which is used to identify the distinctive hydrophobic

portion (h-region) typical of leader peptide sequences. In addition, the presence/absence of a potential ribosomal binding site (Shine-Dalgarno sequence required for translation) is also noted.

2. All putative surface protein sequences are used to search the OWL sequence database which includes a translation of the GENBANK and SWISSPROT database.. This allows identification of similar sequences which may have been previously characterised not only at the sequence level but at a functional level. It may also provide information indicating that these proteins are indeed surface related and not artifacts.

3. Putative *S. agalactiae* surface proteins are also be assessed for their novelty. Some of the identified proteins may or may not possess a typical leader peptide sequence and may not show homology with any DNA/protein sequences in the database. Indeed these proteins may indicate the primary advantage of our screening method, i.e. isolating atypical surface related proteins, which would have been missed in all previously described screening protocols.

The construction of three reporter vectors and their use in *L. lactis* to identify and isolate genomic DNA fragments from pathogenic bacteria encoding secreted or surface associated proteins is now described.

Construction of the pTREP1-*nuc* series of reporter vectors

(a) <u>Construction of expression plasmid pTREP1</u>

The pTREP1 plasmid is a high-copy number (40-80 per cell) theta-replicating gram positive plasmid, which is a derivative of the pTREX plasmid which is itself a derivative of the the previously published pIL253 plasmid. pIL253 incorporates the broad Gram-positive host range replicon of pAMβ1 (Simon and Chopin, 1988) and is non-mobilisable by the *L lactis* sex-factor. pIL253 also lacks the *tra* function which is

necessary for transfer or efficient mobilisation by conjugative parent plasmids exemplified by pIL501. The Enterococcal pAMβ1 replicon has previously been transferred to various species including *Streptococcus, Lactobacillus* and *Bacillus* species as well as *Clostridium acetobutylicum*, (LeBlanc *et al.*, *Proceedings of the National Academy of Science USA* **75**:3484-3487 (1978)) indicating the potential broad host range utility. The pTREP1 plasmid represents a constitutive transcription vector.

The pTREX vector was constructed as follows. An artificial DNA fragment containing a putative RNA stabilising sequence, a translation initiation region (TIR), a multiple cloning site for insertion of the target genes and a transcription terminator was created by annealing 2 complementary oligonucleotides and extending with Tfl DNA polymerase. The sense and anti-sense oligonucleotides contained the recognition sites for NheI and BamHI at their 5' ends respectively to facilitate cloning. This fragment was cloned between the XbaI and BamHI sites in pUC19NT7, a derivative of pUC19 which contains the T7 expression cassette from pLET1 (Wells *et al., J. Appl. Bacteriol.* **74**:629-636 (1993)) cloned between the EcoRI and HindIII sites. The resulting construct was designated pUCLEX. The complete expression cassette of pUCLEX was then removed by cutting with HindIII and blunting followed by cutting with EcoRI before cloning into EcoRI and SacI (blunted) sites of pIL253 to generate the vector pTREX (Wells and Schofield, *In* Current advances in metabolism, genetics and applications-NATO ASI Series. **H 98**:37-62. (1996)). The putative RNA stabilising sequence and TIR are derived from the *Escherichia coli* T7 bacteriophage sequence and modified at one nucleotide position to enhance the complementarity of the Shine Dalgarno (SD) motif to the ribosomal 16s RNA of *Lactococcus lactis* (Schofield *et al.* pers. coms. University of Cambridge Dept. Pathology.).

A *Lactococcus lactis* MG1363 chromosomal DNA fragment exhibiting promoter activity which was subsequently designated P7 was cloned between the EcoRI and

BglII sites present in the expression cassette, creating pTREX7. This active promoter region had been previously isolated using the promoter probe vector pSB292 (Waterfield *et al.*, *Gene* **165**:9-15 (1995)). The promoter fragment was amplified by PCR using the Vent DNA polymerase according to the manufacturer.

The pTREP1 vector was then constructed as follows. An artificial DNA fragment which included a transcription terminator, the forward pUC sequencing primer, a promoter multiple cloning site region and a universal translation stop sequence was created by annealing two overlapping partially complementary synthetic oligonucleotides together and extending with sequenase according to manufacturers instructions. The sense and anti-sense ($pTREP_F$ and $pTREP_R$) oligonucleotides contained the recognition sites for EcoRV and BamHI at their 5' ends respectively to facilitate cloning into pTREX7. The transcription terminator was that of the *Bacillus penicillinase* gene, which has been shown to be effective in Lactococcus (Jos *et al.*, *Applied and Environmental Microbiology* **50**:540-542 (1985)). This was considered necessary as expression of target genes in the pTREX vectors was observed to be leaky and is thought to be the result of cryptic promoter activity in the origin region (Schofield *et al.* pers. coms. University of Cambridge Dept. Pathology.). The forward pUC primer sequencing was included to enable direct sequencing of cloned DNA fragments. The translation stop sequence which encodes a stop codon in 3 different frames was included to prevent translational fusions between vector genes and cloned DNA fragments. The pTREX7 vector was first digested with EcoRI and blunted using the 5' - 3' polymerase activity of T4 DNA polymerase (NEB) according to manufacturer's instructions. The EcoRI digested and blunt ended pTREX7 vector was then digested with Bgl II thus removing the P7 promoter. The artificial DNA fragment derived from the annealed synthetic oligonucleotides was then digested with EcoRV and Bam HI and cloned into the EcoRI(blunted)-Bgl II digested pTREX7 vector to generate pTREP. A *Lactococcus lactis* MG1363 chromosomal promoter designated P1 was then cloned between the EcoRI and BglII sites present in the pTREP expression

cassette forming pTREP1. This promoter was also isolated using the promoter probe vector pSB292 and characterised by Waterfield *et al.*, (1995) [*supra*]. The P1 promoter fragment was originally amplified by PCR using vent DNA polymerase according to manufacturers instructions and cloned into the pTREX as an EcoRI-BglII DNA fragment. The EcoRI-BglII P1 promoter containing fragment was removed from pTREX1 by restriction enzyme digestion and used for cloning into pTREP (Schofield *et al.* pers. coms. University of Cambridge, Dept. Pathology.).

(b) PCR amplification of the *S. aureus nuc* gene.

The nucleotide sequence of the *S. aureus nuc* gene (EMBL database accession number V01281) was used to design synthetic oligonucleotide primers for PCR amplification. The primers were designed to amplify the mature form of the *nuc* gene designated *nuc*A which is generated by proteolytic cleavage of the N-terminal 19 to 21 amino acids of the secreted propeptide designated Snase B (Shortle, 1983 [*supra*]). Three sense primers (*nuc*S1, *nuc*S2 and *nuc*S3, shown in figure 3) were designed, each one having a blunt-ended restriction endonuclease cleavage site for EcoRV or SmaI in a different reading frame with respect to the *nuc* gene. Additionally BglII and BamHI were incorporated at the 5' ends of the sense and anti-sense primers respectively to facilitate cloning into BamHI and BglII cut pTREP1. The sequences of all the primers are given in figure 3. Three *nuc* gene DNA fragments encoding the mature form of the nuclease gene (*Nuc*A) were amplified by PCR using each of the sense primers combined with the anti-sense primer. The *nuc* gene fragments were amplified by PCR using *S. aureus* genomic DNA template, Vent DNA Polymerase (NEB) and the conditions recommended by the manufacturer. An initial denaturation step at 93°C for 2 min was followed by 30 cycles of denaturation at 93°C for 45 sec, annealing at 50°C for 45 seconds, and extension 73°C for 1 minute and then a final 5 min extension step at 73°C. The PCR amplified products were purified using a Wizard clean up column (Promega) to remove unincorporated nucleotides and primers.

(c) Construction of the pTREP1-*nuc* vectors

The purified *nuc* gene fragments described in section b were digested with Bgl II and BamHI using standard conditions and ligated to BamHI and BglII cut and dephosphorylated pTREP1 to generate the pTREP1-*nuc*1, pTREP1-*nuc*2 and pTREP1-*nuc*3 series of reporter vectors. These vectors are described in figure 4. General molecular biology techniques were carried out using the reagents and buffers supplied by the manufacturer or using standard techniques (Sambrook and Maniatis, Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory Press: Cold Spring Harbour (1989)). In each of the pTREP1-*nuc* vectors the expression cassette comprises a transcription terminator, lactococcal promoter P1, unique cloning sites (BglII, EcoRV or SmaI) followed by the mature form of the *nuc* gene and a second transcription terminator. Note that the sequences required for translation and secretion of the *nuc* gene were deliberately excluded in this construction. Such elements can only be provided by appropriately digested foreign DNA fragments (representing the target bacterium) which can be cloned into the unique restriction sites present immediately upstream of the *nuc* gene.

(d) Screening for secreted proteins in *Group B* Streptococcus.

Genomic DNA isolated from and *Group B Streptococcus (S. agalactiae)* was digested with the restriction enzyme Tru9I. This enzyme which recognises the sequence 5'-TTAA -3' was used because it cuts A/T rich genomes efficiently and can generate random genomic DNA fragments within the preferred size range (usually averaging 0.5 - 1.0 kb). This size range was preferred because there is an increased probability that the P1 promoter can be utilised to transcribe a novel gene sequence. However, the P1 promoter may not be necessary in all cases as it is possible that many Streptococcal promoters are recognised in *L. lactis*. DNA fragments of different size ranges were purified from partial Tru9I digests of and *S. agalactiae* genomic DNA. As the Tru 9I restriction enzyme generates staggered ends the DNA fragments had to be made blunt ended before ligation to the EcoRV or SmaI cut pTREP1-*nuc* vectors. This was

achieved by the partial fill-in enzyme reaction using the 5'-3' polymerase activity of Klenow enzyme. Briefly Tru9I digested DNA was dissolved in a solution (usually between 10-20 µl in total) supplemented with T4 DNA ligase buffer (New England Biolabs; NEB) (1X) and 33 µM of each of the required dNTPs, in this case dATP and dTTP. Klenow enzyme was added (1 unit Klenow enzyme (NEB) per µg of DNA) and the reaction incubated at 25°C for 15 minutes. The reaction was stoped by incubating the mix at 75°C for 20 minutes. EcoRV or SmaI digested pTREP-*nuc* plasmid DNA was then added (usually between 200-400 ng). The mix was then supplemented with 400 units of T4 DNA ligase (NEB) and T4 DNA ligase buffer (1X) and incubated overnight at 16°C. The ligation mix was precipiated directly in 100% Ethanol and 1/10 volume of 3M sodium acetate (pH 5.2) and used to transform *L. lactis* MG1363 (Gasson, *J. Bacteriol.* **154:**1-9 (1983)). Alternatively, the gene cloning site of the pTREP-*nuc* vectors also contains a BglII site which can be used to clone for example Sau3AI digested genomic DNA fragments.

L. lactis transformant colonies were grown on brain heart infusion agar and nuclease secreting (*Nuc*$^+$) clones were detected by a toluidine blue-DNA-agar overlay (0.05 M Tris pH 9.0, 10 g of agar per litre, 10 g of NaCl per liter, 0.1 mM CaCl2, 0.03% wt/vol. salmon sperm DNA and 90 mg of Toluidine blue O dye) essentially as described by Shortle, 1983 [*supra*], and Le Loir *et al.*, 1994 [*supra*]). The plates were then incubated at 37°C for up to 2 hours. Nuclease secreting clones develop an easily identifiable pink halo. Plasmid DNA was isolated from *Nuc*$^+$ recombinant *L. lactis* clones and DNA inserts were sequenced on one strand using the *Nuc*Seq sequencing primer described in figure 3, which sequences directly through the DNA insert.

Whilst the example described above related specifically to *Group B Streptococcus*, it will be apparent to one skilled in the art that the same screening technique may be used to detect exported and secreted proteins in other gram positive bacteria, for example *Streptococcus pneumoniae*.

Claims:

1. A *Group B Streptococcus* protein having a sequence selected from those described in fig 1, or fragments or derivatives thereof.

2. A *Group B Streptococcus* polypeptide or peptide having a sequence selected from those described in fig 2, or fragments or derivatives thereof.

3. Derivatives or variants of the proteins, polypeptides, and peptides as claimed in claims 1 and 2 which show at least 50% identity to those proteins, polypeptides and peptides claimed in claims 1 and 2.

4. A nucleic molecule comprising or consisting of a sequence which is:

(i) any of the DNA sequences set out in figure 1 and figure 2 herein or their RNA equivalents;

(ii) a sequence which is complementary to any of the sequences of (i);

(iii) a sequence which codes for the same protein or polypeptide, as those sequences of (i) or (ii);

(iv) a sequence which shows substantial identity with any of those of (i), (ii) and (iii); or

(v) a sequence which codes for a derivative, or fragment of a nucleic acid molecule shown in figure 1 or figure 2.

5. A vector comprising DNA encoding for the expression of any one or more proteins, polypeptides, peptides, fragments or derivitives thereof, as claimed in claims 1 to 3.

6. A vector as claimed in claim 5 further comprising DNA encoding any one or more of the following: promoters, enhancers, signal sequences, leader sequences,

translation start and stop signals, DNA stability controlling regions, or a fusion partner.

7. The use of a vector as claimed in claims 5 and 6 in the transformation or transfection of a prokaryotic or eukaryotic host.

8. A host cell suitable for the transformation of vector as claimed in claims 5 and 6.

9. An antibody, an affibody, or a derivative thereof which binds to one or more of the proteins, polypeptides, peptides, fragments or derivatives thereof, as claimed in any one of claims 1 to 3.

10. An immunogenic composition comprising one or more of the proteins, polypeptides, peptides, fragments or derivatives thereof, or nucleic acid sequences as claimed in any one or more of claims 1-3 and claim 4.

11. An immunogenic composition as claimed in claim 10 which is a vaccine.

12. Use of an immunogenic composition as a claimed in claim 10 in the preparation of a medicament for the treatment or prophylaxis of *Group B Streptococcus* infection.

13. A method of detection of *Group B Streptococcus* which comprises the step of bringing into contact a sample to be tested with at least one antibody, affibody, or a derivative thereof, as described herein.

14. A method of detection of *Group B Streptococcus* which comprises the step of bringing into contact a sample to be tested with at least one protein, polypeptide, peptide, fragments or derivatives as described herein.

15. A method of detection of *Group B Streptococcus* which comprises the step of bringing into contact a sample to be tested with at least one nucleic acid molecule as described herein.

16. A kit for the detection of *Group B Streptococcus* comprising at least one antibody, affibody, or derivatives thereof as claimed in claim 9.

17. A kit for the detection of *Group B Streptococcus* comprising at least one *Group B Streptococcus* protein, polypeptide, peptide, fragment or derivative thereof as claimed in claims 1 to 3.

18. A kit for the detection of *Group B Streptococcus* comprising at least one nucleic acid molecule as claimed in claim 4.

19. A method of screening for DNA encoding bacterial cell envelope associated or surface antigens in gram positive bacteria comprising the steps of:

- combining a reporter vector including the nucleotide sequence encoding the mature from of the staphylcoccus nuclease gene and an upstream promoter region with DNA from a gram positive bacteria.
- transforming the resultant vector into *Lactococcus lactis* cells.
- assaying for the secretion of staphlycoccus nuclease protein in the transformed cells.

20. A method as claimed in claim 19 wherein the reporter vector is one of the pTREP1-*nuc* vectors shown in figure 4.

21. A method as claimed in claim 19 or claim 20 wherein the gram positive bacteria is *Group B Streptococcus, Streptococcus Pneumoniae, Staphylcoccus aureus* or *pathogenic group A streptococci.*

22. A vector as shown in figure 4 for use in screening for DNA encoding bacterial cell envelope associated or secreted antigens in gram positive bacteria.

23. A method of determining whether a protein, polypeptide, peptide, fragment or derivative thereof as claimed in claims 1 to 3 represents a potential anti-microbial target which comprises inactivating said protein and determining whether *Group B Streptococcus* is still viable.

FIGURE 1

ID-1: 1248 base pairs
Clone 4
(A)
ATGGAAAAAAATACTTGGAAAAAATTACTTGTTAGTACTGCTGCTCTTTCAGTAGTTGCAGGA
GGAGCAATTGCTGCTACTCACTCTAACTCAGTTGATGCTGCTTCAAAAAAAACTATCAAACTT
TGGGTCCCAACAGATTCAAAAGCGTCTTATAAAGCAATTGTTAAAAAATTCGAGAAGGAAAAC
AAAGGCGTTACTGTAAAAATGATTGAGTCTAATGACTCCAAAGCTCAAGAAAACGTAAAAAAA
GACCCAAGCAAGGCAGCCGATGTATTCTCACTTCCACATGACCAACTTGGTCAATTAGTAGAA
TCTGGTGTTATCCAAGAAATTCCAGAGCAATACTCAAAAGAAATTGCTAAAAACGACACTAAA
CAATCACTTACTGGTGCACAATATAAAGGGAAAACTTATGCATTCCCATTTGGTATTGAATCT
CAAGTTCTTTATTATAATAAAACAAAGTTAACTGCTGACGACGTTAAATCATACGAAACAATT
ACAAGCAAAGGGAAATTCGGTCAACAGCTTAAAGCAGCTAACTCATATGTAACAGGTCCTCTT
TTCCTTTCTGTAGGCGACACTTTATTTGGTAAATCTGGTGAAGATGCTAAAGGCACTAACTGG
GGTAATGAAGCAGGTGTTTCTGTCCTTAAATGGATTGCAGATCAAAAGAAAAATGATGGTTTT
GTCAACTTGACAGCTGAAAATACAATGTCTAAATTTGGCGATGGTTCTGTTCATGCTTTTGAA
AGTGGACCATGGGATTACGACGCTGCTAAAAAAGCTGTCGGTGAAGATAAAATCGGTGTTGCT
GTTTACCCAACAATGAAAATCGGTGACAAAGAAGTTCAACAAAAAGCATTCTTGGGCGTTAAA
CTTTATGCCGTTAACCAAGCACCTGCTGGTTCAAACACTAAACGAATCTCAGCTAGCTACAAA
CTCGCTGCATATCTAACTAATGCTGAAAGTCAAAAAATTCAATTCGAAAAACGTCATATCGTT
CCTGCTAACTCATCAATTCAATCTTCTGATAGCGTCCAAAAAGATGAACTTGCAAAAGCAGTT
ATCGAAATGGGTAGCTCAGATAAATATACAACGGTTATGCCTAAGTTGAGTCAAATGTCAACA
TTCTGGACAGAAAGTGCTGCTATTCTTAGCGATACTTACAGTGGTAAAATCAAATCTAGCGAT
TACCTTAAACGTCTAAAACAATTCGATAAAGACATCGCTAAAACAAAATAG

(B)
MEKNTWKKLLVSTAALSVVAGGAIAATHSNSVDAASKKTIKLWVPTDSKASYKAIVKKFEKEN
KGVTVKMIESNDSKAQENVKKDPSKAADVFSLPHDQLGQLVESGVIQEIPEQYSKEIAKNDTK
QSLTGAQYKGKTYAFPFGIESQVLYYNKTKLTADDVKSYETITSKGKFGQQLKAANSYVTGPL
FLSVGDTLFGKSGEDAKGTNWGNEAGVSVLKWIADQKKNDGFVNLTAENTMSKFGDGSVHAFE
SGPWDYDAAKKAVGEDKIGVAVYPTMKIGDKEVQQKAFLGVKLYAVNQAPAGSNTKRISASYK
LAAYLTNAESQKIQFEKRHIVPANSSIQSSDSVQKDELAKAVIEMGSSDKYTTVMPKLSQMST
FWTESAAILSDTYSGKIKSSDYLKRLKQFDKDIAKTKZ

ID-2: 1539 base pairs
Clone 5
(A)
ATGTCAAAACAAAAAGTAACGGCAACTTTGTTGTTATCCACTTTAGTCTTATCGCTATCATCA
CCTTTAGTGACCTTAGCAGAAACTATTAATCCAGAAACAAGCCTGACAATGGCAACAGCATCA
ACAGAAAGTTCTTCTGAAGCAGAGAAACAGGAAAAAACACAACCTACAGATTCAGAAACTGCT
TCACCTTCAGCCGAAGGAAGTATCTCAACAGAAAAAACAGAGATTGGTACGACAGAGACATCA

TCAAGCAATGAATCATCATCAAGTTCATCACATCAATCTTCTTCCAACGAAGATGCTAAAACA
TCTGATTCTGCTTCAACAGCATCTACTCCTAGCACTAATACTACAAACAGTAGTCAAGCAGAC
AGTAAGCCAGGTCAATCAACAAAGACTGAATTAAAACCTGAGCCTACCTTACCATTAGTAGAG
CCTAAAATAACTCCCGCTCCGTCTCAGATAGAAAGTGTTCAGACAAATCAGAATGCTTCTGTT
CCTGCTTTATCCTTTGATGATAACTTATTATCAACACCGATTTCACCAGTGACAGCAACGCCA
TTCTACGTAGAACACTGGTCTGGTCAGGATGCCTACTCTCACTATTTATTGTCACATCGTTAC
GGTATCAAAGCTGAACAATTAGATGGGTACTTAAAATCTTTAGGGATTCAATATGATTCTAAT
CGTATCAATGGTGCTAAGTTATTACAATGGGAAAAAGATAGTGGTTTAGATGTCCGTGCTATT
GTAGCTATTGCTGTCCTTGAAAGTTCATTGGGAACTCAAGGAGTGGCTAAAATGCCAGGTGCT
AATATGTTTGGTTATGGTGCCTTTGATCATGACTCTAGCCATGCTAGTGCTTATAATGATGAA
GAAGCAATTATGTTGTTGACAAAAAATACAATTATTAAAAACAACAACTCTAGCTTTGAAATC
CAAGATTTGAAAGCACAGAAATTATCTTCTGGACAACTTAATACAGTTACTGAGGGTGGTGTT
TATTATACAGATAACTCTGGAACTGGTAAACGTCGTGCCCAGATTATGGAAGATTTAGACCGC
TGGATTGATCAACATGGAGGGACACCAGAAATTCCTGCTGCCTTGAAAGCTTTATCGACAGCA
AGTTTAGCAGATTTACCAAGTGGTTTTAGCTTATCAACAGCGGTTAACACAGCTAGCTATATT
GCATCAACTTATCCATGGGGTGAATGTACATGGTATGTCTTTAACCGCGCTAAAGAGTTAGGT
TATACATTTGATCCATTTATGGGTAATGGTGGAGATTGGCAACATAAGGCTGGCTTTGAAACA
ACACATTCACCAAAAGTAGGCTATGCTGTATCATTTTCACCAGGACAAGCTGGTGCTGATGGC
ACTTACGGTCACGTAGCTATTGTTGAAGAAGTTAAAAAAGATGGTTCAGTTCTCATTTCAGAA
TCTAATGCAATGGGACGTGGTATTGTCTCTTACCGTACTTTTAGTTCAGCACAAGCTGCACAA
TTAACTTATGGTATTGGCCATAAATAA

(B)
MSKQKVTATLLLSTLVLSLSSPLVTLAETINPETSLTMATASTESSSEAEKQEKTQPTDSETA
SPSAEGSISTEKTEIGTTETSSSNESSSSSSHQSSSNEDAKTSDSASTASTPSTNTTNSSQAD
SKPGQSTKTELKPEPTLPLVEPKITPAPSQIESVQTNQNASVPALSFDDNLLSTPISPVTATP
FYVEHWSGQDAYSHYLLSHRYGIKAEQLDGYLKSLGIQYDSNRINGAKLLQWEKDSGLDVRAI
VAIAVLESSLGTQGVAKMPGANMFGYGAFDHDSSHASAYNDEEAIMLLTKNTIIKNNNSSFEI
QDLKAQKLSSGQLNTVTEGGVYYTDNSGTGKRRAQIMEDLDRWIDQHGGTPEIPAALKALSTA
SLADLPSGFSLSTAVNTASYIASTYPWGECTWYVFNRAKELGYTFDPFMGNGGDWQHKAGFET
THSPKVGYAVSFSPGQAGADGTYGHVAIVEEVKKDGSVLISESNAMGRGIVSYRTFSSAQAAQ
LTYGIGHKZ

ID-3: 1293 base pairs
Clone 6
(A)
GTGCATATGTTACAAAACATTGGACAAACAGGCATTCAAGCAACTCGAATTGCTTTAGGTTGT
ATGAGAATGAGTGACTTGAAAGGAAAACAAGCTGAAGAAGTAGTTGGAACAGCATTAGATTTG
GGTATTATAAATAATAAAGTGCAAGAAAGTGTCTCTGGCGTCAAAGTGACTAAATCATTGTGT
TATCAAGAACAAGAAATTGCTTCTTTTCAAGAGATTAATCAGATGACTTTCGTGAAGAACATG
CGGACCATGACTTATGATGTCATGTTTGATCCTTTAGTTCTTCTTTTTATAGGTGCCTCCTAC
GTATTAACATTGGCTATGGGAGCTTTTATGATTTCAAAAGGTCAAGTTACTGTTGGTGACTTG
GTAACATTTGTGACGTATTTAGATATGTTGGTATGGCCCTTGATGGCGATTGGTTTCTTGTTC
AATATGGTACAGCGTGGTAGTGTTTCTTATAACCGTATTAATAGTCTACTTGAGCAAGAATCG

GATATAACTGATCCTTTAAATCCTATCAAACCTGTTGTCAATGGAACATTAAGATATGATATT
GATTTCTTTAGATACGACAATGAGGAAACCTTAGCCGATATTCATTTCACCTTAGAAAAAGGT
CAAACCTTAGGTTTGGTAGGTCAAACGGGATCAGGGAAGACAAGTCTTATTAAGTTATTGCTA
CGTGAACATGATGTGACTCAGGGGAAAATTACTTTAAATAAACATGATATACGTGATTATCGA
TTGTCTGAGTTACGTCAACTAATCGGTTATGTTCCTCAAGATCAGTTTTTATTTGCTACCAGT
ATTTTAGAAAATGTTCGCTTTGGAAATCCAACTCTATCTATCAATGCTGTCAAAGAAGCAACT
AAATTGGCACATGTTTACGATGACATTGAACAGATGCCAGCAGGATTTGAGACTCTAATTGGA
GAAAAAGGAGTCTCATTATCTGGTGGACAAAAACAAAGGATTGCGATGAGTCGTGCCATGATT
TTAGATCCAGATATTCTTATTTTGGATGATTCTCTATCAGCAGTGGACGCTAAAACGGAACAT
GCTATTGTTGAGAATCTTAAAACGAATCGTCAAGGGAAATCGACTATTATTTCAGCACATCGT
TTATCAGCTGTTGTGCACGCAGACCTTATCTTAGTTATGCGAGACGGCAGAGTCATTGAGCGA
GGTCAACATCAAGAGTTGCTAAATAAAGGTGGTTGGTATGCTGAAACGTATGCCTCACAGCAA
TTAGAAATGGAGGAAGCATTTGATGAAGTCTAA

(B)
MHMLQNIGQTGIQATRIALGCMRMSDLKGKQAEEVVGTALDLGIINNKVQESVSGVKVTKSLC
YQEQEIASFQEINQMTFVKNMRTMTYDVMFDPLVLLFIGASYVLTLAMGAFMISKGQVTVGDL
VTFVTYLDMLVWPLMAIGFLFNMVQRGSVSYNRINSLLEQESDITDPLNPIKPVVNGTLRYDI
DFFRYDNEETLADIHFTLEKGQTLGLVGQTGSGKTSLIKLLLREHDVTQGKITLNKHDIRDYR
LSELRQLIGYVPQDQFLFATSILENVRFGNPTLSINAVKEATKLAHVYDDIEQMPAGFETLIG
EKGVSLSGGQKQRIAMSRAMILDPDILILDDSLSAVDAKTEHAIVENLKTNRQGKSTIISAHR
LSAVVHADLILVMRDGRVIERGQHQELLNKGGWYAETYASQQLEMEEAFDEVZ

ID-6: 921 base pairs
Clone 9
(A)
ATGAAAAAAGTTTTTTTTCTCATGGCTATGGTTGTGAGTTTAGTAATGATAGCAGGGTGTGAT
AAGTCAGCAAACCCCAAACAGCCTACGCAAGGCATGTCAGTTGTAACCAGCTTTTACCCAATG
TATGCGATGACAAAAGAAGTATCTGGAGACCTAAATGATGTGAGGATGATCCAATCAGGTGCA
GGCATTCATTCCTTTGAACCGTCTGTAAATGATGTGGCAGCTATTTATGACGCGGATTTGTTT
GTTTACCAATCACATACCTTAGAAGCTTGGGCAAGGGATCTAGACCCTAATTTAAAAAAATCA
AAGGTTAATGTGTTTGAAGCGTCAAAACCTCTGACACTAGATAGAGTCAAAGGGCTAGAAGAT
ATGGAAGTCACACAAGGCATTGACCCTGCGACACTTTATGACCCACATACCTGGACGGATCCC
GTTTTAGCTGGTGAGGAAGCTGTTAATATCGCTAAAGAGCTAGGACATTTGGATCCTAAACAC
AAAGACAGTTACACTAAAAAGGCTAAGGCTTTCAAAAAAGAAGCAGAGCAACTAACTGAAGAA
TACACTCAAAAATTTAAAAAGGTGCGCTCAAAAACATTTGTGACGCAACACACGGCATTTTCT
TATCTGGCTAAACGATTCGGCTTGAAACAACTTGGTATCTCGGGTATTTCTCCAGAGCAAGAG
CCCTCTCCTCGCCAATTGAAAGAAATTCAAGACTTTGTTAAAGAATACAACGTCAAGACTATT
TTTGCAGAAGACAACGTCAACCCCAAAATTGCTCATGCTATTGCGAAATCAACAGGAGCTAAA
GTAAAGACATTAAGTCCACTTGAAGCTGCTCCAAGCGGAAACAAGACATATCTAGAAAATCTT
AGAGCAAATTTGGAAGTGCTCTATCAACAGTTGAAGTAA

(B)
MKKVFFLMAMVVSLVMIAGCDKSANPKQPTQGMSVVTSFYPMYAMTKEVSGDLNDVRMIQSGA
GIHSFEPSVNDVAAIYDADLFVYQSHTLEAWARDLDPNLKKSKVNVFEASKPLTLDRVKGLED
MEVTQGIDPATLYDPHTWTDPVLAGEEAVNIAKELGHLDPKHKDSYTKKAKAFKKEAEQLTEE
YTQKFKKVRSKTFVTQHTAFSYLAKRFGLKQLGISGISPEQEPSPRQLKEIQDFVKEYNVKTI
FAEDNVNPKIAHAIAKSTGAKVKTLSPLEAAPSGNKTYLENLRANLEVLYQQLKZ

ID-8:1029 base pairs
Clone 17
(A)
ATGACAAAAAAACTTATTATTGCTATATTAGCACTATGCACTATCTTAACCACTTCTCAAGCT
GTTTTAGCTAAAGAAAAATCACAAACTGTTACCATAAAAAACAACTATTCGGTCTATATTAAA
AAAGAAAAAAGAGACAAGCCGGATAATAAAAAGCAAATCAGCGAGACACTTAAAGTTCCTTTA
AAACCCAAAAAAGTAGTTGTTTTTGATATGGGAGCTTTGGATACTATCACAGCTTTAGGAGCT
GAAAAATCTGTTATTGGTATCCCGAAGGCTAAAAATGCTCTAAGTTTATTGCCCAATAACGTC
AAATCTGTTTATAAAGCTAAGAGATACCAAGACGTAGGAAGTCTCTTCGAACCAAACTTTGAA
GCTATTGCTCGTATGCAACCTGATGTGGTTTTCCTAGGAGCACGTATGGCTTCTGTTGATAAT
ATTGAAAAATTAAAGGAGGCTGCACCTAAAGCAGCATTAGTATATGCTGGAGTCGACTCAAAA
AAAGTATTTGACAAAGGAGTTGCTGAGCGTGTCACAATGTTAGGGAAAATCTTCGACCAAAAT
AAAAAGGCAAAAACCTTTAATAAAGATATCGCACAAGCTGTTCTTAAATTGCAGAAAACTATT
GAGAAAAAAGGTAAACCTACAGCTCTATTTGTAATGGCAAACAGCGGTGAACTTTTAACTCAA
TCACCTTCTGGTCGTTTTGGTTGGATTTTCTCTGTAGGTGGATTTAAAGCAGTCAATGAAAAT
GAAAAACTAAGTTCACATGGTACTCCCGTATCTTATGAATACATCGCTGAAAAAAATCCTAAC
TATCTCTTTGTTTTAGATCGTGGAGCGACTATTGGACAAGGAGCTTCATCAAAAGAACTTTTT
AATAACGATGTTATTAAAGCAACTGATGCTGTCAAAAACAAACGTGTTCATGAGGTAGATGGA
AAAGATTGGTATATCAATTCAGGCGGAAGCCGAGTAACACTCCGTATGATTAAAGATGTACAG
AACTTTGTTGATAATCGTTAA

(B)
MTKKLIIAILALCTILTTSQAVLAKEKSQTVTIKNNYSVYIKKEKRDKPDNKKQISETLKVPL
KPKKVVVFDMGALDTITALGAEKSVIGIPKAKNALSLLPNNVKSVYKAKRYQDVGSLFEPNFE
AIARMQPDVVFLGARMASVDNIEKLKEAAPKAALVYAGVDSKKVFDKGVAERVTMLGKIFDQN
KKAKTFNKDIAQAVLKLQKTIEKKGKPTALFVMANSGELLTQSPSGRFGWIFSVGGFKAVNEN
EKLSSHGTPVSYEYIAEKNPNYLFVLDRGATIGQGASSKELFNNDVIKATDAVKNKRVHEVDG
KDWYINSGGSRVTLRMIKDVQNFVDNRZ

ID-9: 2469 base pairs
Clone 18
(A)
GTGAAGAAAACATATGGTTATATCGGCTCAGTTGCTGCTATTTTACTAGCTACTCATATTGGA
AGTTACCAGCTTGGTAAGCATCATATGGGTCTAGCAACAAAGGACAATCAGATTGCCTATATT
GATGATAGCAAAGGTAAGGTAAAAGCCCCTAAAACAAACAAAACGATGGATCAAATCAGTGCT
GAAGAAGGCATCTCTGCTGAACAGATCGTAGTCAAAATTACTGACCAAGGTTATGTTACCTCA

CACGGTGACCATTATCATTTTTACAATGGGAAAGTTCCTTATGATGCGATTATTAGTGAAGAG
TTGTTGATGACGGATCCTAATTACCATTTTAAACAATCAGACGTTATCAATGAAATCTTAGAC
GGTTACGTTATTAAAGTCAATGGCAACTATTATGTTTACCTCAAGCCAGGTAGTAAGCGCAAA
AACATTCGAACCAAACAACAAATTGCTGAGCAAGTAGCCAAAGGAACTAAAGAAGCTAAAGAA
AAAGGTTTAGCTCAAGTGGCCCATCTCAGTAAAGAAGAAGTTGCGGCAGTCAATGAAGCAAAA
AGACAAGGACGCTATACTACAGACGATGGCTATATTTTTAGTCCGACAGATATCATTGATGAT
TTAGGAGATGCTTATTTAGTACCTCATGGTAATCACTATCATTATATTCCTAAAAAAGATTTG
TCTCCAAGTGAGCTAGCTGCTGCACAAGCCTACTGGAGTCAAAAACAAGGTCGAGGTGCTAGA
CCGTCTGATTACCGCCCGACACCAGCCCCAGGTCGTAGGAAAGCCCCAATTCCTGATGTGACG
CCTAACCCTGGACAAGGTCATCAGCCAGATAACGGTGGTTATCATCCAGCGCCTCCTAGGCCA
AATGATGCGTCACAAAACAAACACCAAAGAGATGAGTTTAAAGGAAAAACCTTTAAGGAACTT
TTAGATCATCTACACCGTCTTGATTTGAAATACCGTCATGTGGAAGAAGATGGGTTGATTTTT
GAACCGACTCAAGTGATCAAATCAAACGCTTTTGGGTATGTGGTGCCTCATGGAGATCATTAT
CATATTATCCCAAGAAGTCAGTTATCACCTCTTGAAATGGAATTAGCAGATCGATACTTAGCC
GGCCAAACTGATGACAACGACTCAGGTTCAGATCACTCAAAACCATCAGATAAAGAAGTGACA
CATACCTTTCTTGGTCATCGCATCAAAGCTTACGGAAAAGGCTTAGATGGTAAACCATATGAT
ACGAGTGATGCTTATGTTTTTAGTAAAGAATCCATTCATTCAGTGGATAAATCAGGAGTTACA
GCTAAACACGGAGATCATTTCCACTATATAGGATTTGGAGAACTTGAACAATATGAGTTGGAT
GAGGTCGCTAACTGGGTGAAAGCAAAAGGTCAAGCTGATGAGCTTGTTGCTGCTTTGGATCAG
GAACAAGGCAAAGAAAAACCACTCTTTGACACTAAAAAAGTGAGTCGCAAAGTAACAAAAGAT
GGTAAAGTGGGCTATATTATGCCAAAAGATGGCAAGGACTATTTCTATGCTCGTTATCAACTT
GATTTGACTCAGATTGCCTTTGCCGAACAAGAACTAATGCTTAAAGATAAGAAGCATTACCGT
TATGACATTGTTGATACAGGCATTGAGCCACGACTTGCTGTAGATGTGTCAAGTCTGCCGATG
CATGCTGGTAATGCTACTTACGATACTGGAAGTTCGTTTGTTATCCCACATATTGATCATATC
CATGTCGTTCCGTATTCATGGTTGACGCGCAATCAGATTGCAACAATCAAGTATGTGATGCAA
CACCCCGAAGTTCGTCCGGATGTATGGTCTAAGCCAGGGCATGAAGAGTCAGGTTCGGTCATT
CCAAATGTTACGCCTCTTGATAAACGTGCTGGTATGCCAAACTGGCAAATTATCCATTCTGCT
GAAGAAGTTCAAAAAGCCCTAGCAGAAGGTCGTTTTGCAGCACCAGACGGCTATATTTTCGAT
CCACGAGATGTTTTGGCAAAAGAAACTTTTGTATGGAAAGATGGCTCCTTTAGCATCCCAAGA
GCAGATGGCAGTTCATTGAGAACCATTAATAAATCCGATCTATCCCAAGCTGAGTGGCAACAA
GCTCAAGAGTTATTGGCAAAGAAAAATGCTGGTGATGCTACTGATACGGATAAACCTGAAGAA
AAGCAACAGGCAGATAAGAGCAATGAAAACCAACAGCCAAGTGAAGCCAGTAAAGAAGAAAAA
GAATCAGATGACTTTATAGACAGTTTACCAGACTATGGTCTAGATAGAGCAACCCTAGAAGAT
CATATCAATCAATTAGCACAAAAAGCTAATATCGATCCTAAGTATCTCATTTTCCAACCAGAA
GGTGTCCAATTTTATAATAAAAATGGTGAATTGGTAACTTATGATATCAAGACACTTCAACAA
ATAAACCCTTAA

(B)

MKKTYGYIGSVAAILLATHIGSYQLGKHHMGLATKDNQIAYIDDSKGKVKAPKTNKTMDQISA
EEGISAEQIVVKITDQGYVTSHGDHYHFYNGKVPYDAIISEELLMTDPNYHFKQSDVINEILD
GYVIKVNGNYYVYLKPGSKRKNIRTKQQIAEQVAKGTKEAKEKGLAQVAHLSKEEVAAVNEAK
RQGRYTTDDGYIFSPTDIIDDLGDAYLVPHGNHYHYIPKKDLSPSELAAAQAYWSQKQGRGAR
PSDYRPTPAPGRRKAPIPDVTPNPGQGHQPDNGGYHPAPPRPNDASQNKHQRDEFKGKTFKEL
LDHLHRLDLKYRHVEEDGLIFEPTQVIKSNAFGYVVPHGDHYHIIPRSQLSPLEMELADRYLA
GQTDDNDSGSDHSKPSDKEVTHTFLGHRIKAYGKGLDGKPYDTSDAYVFSKESIHSVDKSGVT

This Page Blank (uspto)

AKHGDHFHYIGFGELEQYELDEVANWVKAKGQADELVAALDQEQGKEKPLFDTKKVSRKVTKD
GKVGYIMPKDGKDYFYARYQLDLTQIAFAEQELMLKDKKHYRYDIVDTGIEPRLAVDVSSLPM
HAGNATYDTGSSFVIPHIDHIHVVPYSWLTRNQIATIKYVMQHPEVRPDVWSKPGHEESGSVI
PNVTPLDKRAGMPNWQIIHSAEEVQKALAEGRFAAPDGYIFDPRDVLAKETFVWKDGSFSIPR
ADGSSLRTINKSDLSQAEWQQAQELLAKKNAGDATDTDKPEEKQQADKSNENQQPSEASKEEK
ESDDFIDSLPDYGLDRATLEDHINQLAQKANIDPKYLIFQPEGVQFYNKNGELVTYDIKTLQQ
INPZ

ID-10: 939 base pairs
Clone 22
(A)
ATGATACGCCAGTTTTTAAGAGAACACTTGATTTGGTATATTTTATATATCATGATGTTTGTC
CTATTTTTTATTAGTTTCTATCTATATCATTTACCAATGCCCTATTTGTTTAATTCCTTAGGT
TTAAATGTTATTGTTTTACTAGGAATTAGTATTTGGCAATACAGTCGTTACAGGAAAAAAATG
TTACATCTCAAATATTTTAATAGTAGTCAGGACCCCTCTTTCGAACTTCAACCGAGTGATTAC
GCTTATTTTAATATTATTACACAATTAGAAGCTAGAGAAGCGCAAAAAGTTTCTGAAACAATT
GAACAAACCAATCATGTTGCACTTATGATAAAGATGTGGTCGCACCAAATGAAAGTTCCATTG
GCAGCTATTTCATTAATGGCCCAGACAAATCATCTCGATCCTAAGGAAGTTGAACAACAATTA
TTGAAATTGCAACATTATCTTGAAACGTTGTTAGCATTTTTGAAATTTAGACAATATCGTGAC
GATTTTCGTTTTGAAGCTGTTAGCCTTAGAGAAGTAGTAGTAGAAATTATAAAATCGTATAAG
GTTATTTGTCTATCCAAAAGCTTATCTATCATAATTGAAGGCGATAATATCTGGAAAACAGAC
AAAAAGTGGTTAACTTTTGCTCTTTCACAGGTGCTAGATAATGCCATAAAATATTCTAATCCT
GAGTCAAAGATAATAATAAGCATAGGAGAAGAGAGTATTAGAATACAAGACTACGGTATCGGC
ATACTCGAAGAGGATATCCCTAGACTTTTTGAAGATGGCTTTACGGGTTACAACGGTCATGAG
CACCAAAAGGCAACAGGCATGGGGTTATATATGACAAAAGAAGTCTTATCTAGTCTGAATTTG
TCCATTTCGGTGGATAGCAAAATTAATTATGGGACTGCTGTTTCTATACATAAATAA

(B)
MIRQFLREHLIWYILYIMMFVLFFISFYLYHLPMPYLFNSLGLNVIVLLGISIWQYSRYRKKM
LHLKYFNSSQDPSFELQPSDYAYFNIITQLEAREAQKVSETIEQTNHVALMIKMWSHQMKVPL
AAISLMAQTNHLDPKEVEQQLLKLQHYLETLLAFLKFRQYRDDFRFEAVSLREVVVEIIKSYK
VICLSKSLSIIIEGDNIWKTDKKWLTFALSQVLDNAIKYSNPESKIIISIGEESIRIQDYGIG
ILEEDIPRLFEDGFTGYNGHEHQKATGMGLYMTKEVLSSLNLSISVDSKINYGTAVSIHKZ

ID-13: 660 base pairs
Clone 28
(A)
ATGGTAAATGATATATTAGAAAGAATGTATAAAGAGAATATTCCAAAATCTTACCTTACATCC
GTCCCATTAGTTATTTCTCAAAAAGGAAGAACAACCTATTCGTTTAGTATGACTGGTGGTCAA
CAAATAGATGGAGTGAAATTCACACAGATATATGAGGACTATATGAAATTACTCAGTCAAGGT
AAGGATATCGCAGAGTTATATCAAAAATATTCTAAAGAAGAGTTGGCAAATCTAGGCATTAAT
ATTTATCAATCCAATGATATAGAAAGGACTGAGGAAAGAACTTTTGATGAAATTATCAGTTGG
GTTTCCAACCCTTATGCAACAAGACCAATTCAAGAAAGGCACACTATTCAATTAGAGCCAACA

AGATTTTCACTAGAGGATAAGAAAAGAATTGAAGAAGCTGCAGCTCAAGGACTAAGCGAAATC
GACCTTATTGATTTAGTTGACCTATATGATATTAATTTAGACAATACAAGCGTCAATCGCCAT
ATTGTGGGGTTATTGACTAATAACACCCAAGTAACATACTATTTCCAAGAACAATTAAATAAG
GAGTTGCTGTCAATGGCTCACGCTTTAGATAACGTACAACAGGCCTTTATTAAATTATTAAGT
GAAGAGGAGATACGAAAATTTGCTCTTTAA

(B)
MVNDILERMYKENIPKSYLTSVPLVISQKGRTTYSFSMTGGQQIDGVKFTQIYEDYMKLLSQG
KDIAELYQKYSKEELANLGINIYQSNDIERTEERTFDEIISWVSNPYATRPIQERHTIQLEPT
RFSLEDKKRIEEAAAQGLSEIDLIDLVDLYDINLDNTSVNRHIVGLLTNNTQVTYYFQEQLNK
ELLSMAHALDNVQQAFIKLLSEEEIRKFALZ

ID-14: 654 base pairs
Clone 31
(A)
ATGAATAAAAGAAGAAAATTATCAAAATTGAATGTAAAAAAACAACATTTAGCTTATGGAGCT
ATCACTTTAGTAGCCCTTTTTTCATGTATTTTGGCTGTAACGGTCATCTTTAAAAGTTCACAA
GTTACTACTGAATCTTTGTCAAAAGCAGATAAAGTTCGCGTAGCCAAAAAATCAAAAATGACT
AAGGCGACATCTAAATCAAAAGTAGAAGATGTAAAACAGGCTCCAAAACCTTCTCAGGCATCT
AATGAAGCCCCAAAATCAAGTTCTCAATCTACAGAAGCTAATTCTCAGCAACAAGTTACTGCG
AGTGAAGAGGCGGCTGTAGAACAAGCAGTTGTAACAGAAAATACCCCTGCTACCAGTCAGGCA
CAACAAACTTATGCTGTTACTGAGACAACTTACAAACCTGCTCAACACCAGACAAGTGGCCAA
GTATTGAGCAATGGAAATACTGCAGGGGCGGTCGGATCTGCTGCTGCAGCACAAATGGCTGCT
GCAACAGGAGTCCCTCAGTCTACTTGGGAACATATTATTGCCCGTGAATCAAATGGTAATCCT
AATGTTGCTAATGCCTCAGGGAGCTTCAGGACTTTTCCAAACGATGCCAGGTTGGGGTTCAAC
AGCTACAGTTCAGGATCAAGTTAA

(B)
MNKRRKLSKLNVKKQHLAYGAITLVALFSCILAVTVIFKSSQVTTESLSKADKVRVAKKSKMT
KATSKSKVEDVKQAPKPSQASNEAPKSSSQSTEANSQQQVTASEEAAVEQAVVTENTPATSQA
QQTYAVTETTYKPAQHQTSGQVLSNGNTAGAVGSAAAAQMAAATGVPQSTWEHIIARESNGNP
NVANASGSFRTFPNDARLGFNSYSSGSSZ

ID-15: 360 base pairs
Clone 32
(A)
ATGATTGTTGGACACGGAATTGATTTACAAGAGATAGAGGCGATTACTAAAGCATATGAGCGT
AATCAACGTTTTGCAGAACGCGTTTTGACCGAACAAGAATTGCTTCTTTTTAAAGGAATTTCC
AATCCCAAGCGTCAGATGTCTTTTTTAACAGGGCGATGGGCAGCAAAAGAGGCTTATAGCAAA
GCACTTGGAACAGGAATTGGGAAAGTTAATTTTCATGATATCGAAATTTTATCGGATGATAAA
GGAGCGCCTTTGATTACAAAAGAACCGTTTAATGGAAAATCTTTTGTTTCAATATCTCATAGT
GGTAATTATGCACAAGCTAGTGTTATTTTGGAGGAAGAAAAATGA

(B)
MIVGHGIDLQEIEAITKAYERNQRFAERVLTEQELLLFKGISNPKRQMSFLTGRWAAKEAYSK
ALGTGIGKVNFHDIEILSDDKGAPLITKEPFNGKSFVSISHSGNYAQASVILEEEKZ

ID-16: 474 base pairs
Clone 35
(A)
ATGATTTTTGTCACAGTGGGGACACATGAACAGCAGTTCAACCGTCTTATTAAAGAAGTTGAT
AGATTAAAAGGGACAGGTGCTATTGATCAAGAAGTGTTCATTCAAACGGGTTACTCAGACTTC
GAACCTCAGAATTGTCAGTGGTCAAAATTTCTCTCATATGATGATATGAACTCTTACATGAAA
GAAGCTGAGATTGTTATCACACATGGCGGCCCAGCGACGTTTATGTCAGTTATTTCTTTAGGG
AAATTACCAGTTGTTGTTCCTAGGAGAAAGCAGTTTGGTGAACATATCAATGATCATCAAATA
CAATTTTTAAAAAAAATTGCCCACCTGTATCCCTTGGCTTGGATTGAAGATGTAGATGGACTT
GCGGAAGCGTTGAAAAGGAATATAGCTACAGAAAAATATCAGGGAAATAATGATATGTTTTGT
CATAAATTAGAAAAAATTATAGGTGAAATATGA

(B)
MIFVTVGTHEQQFNRLIKEVDRLKGTGAIDQEVFIQTGYSDFEPQNCQWSKFLSYDDMNSYMK
EAEIVITHGGPATFMSVISLGKLPVVVPRRKQFGEHINDHQIQFLKKIAHLYPLAWIEDVDGL
AEALKRNIATEKYQGNNDMFCHKLEKIIGEIZ

ID-17: 1203 base pairs
Clone39
(A)
TTGGAAGACAAATTATTCAACAAACATTTTATAGGCATTACTATTTTAAACTTTATTGTTTAT
ATGGTCTATTATTTGTTCACCGTTATCATAGCTTTTATTGCGACTAAAGAGTTAGGTGTTAGC
ACTAGCCAAGCAGGATTAGCAACGGGGATTTATATTGTAGGGACTTTGATTGCTCGTCTTATA
TTTGGTAAGCAATTAGAAGTTCTAGGACGTAAGTTAGTTTTACGTGGAGGGGCTATTTTTTAC
TTACTAACAACTTTAGCTTATTTTTATATGCCAAGTATCGGAGTAATGTATTTAGTTCGTTTC
CTAAATGGTTTTGGTTATGGCGTCGTGTCAACAGCAACTAATACTATTGTAACAGCCTATATA
CCAGCTGATAAAAGAGGTGAGGGGATTAACTTTTACGGTCTATCAACAAGTTTAGCCGCAGCT
ATTGGTCCTTTTGTAGGAACATTTATGCTAGACAACCTTCATATTAACTTTAAAATGGTTATT
GTATTATGTAGTATTTTAATTGCGATTGTAGTGTTGGGAGCATTTGTTTTCCCAGTCAAAAAT
ATTACTTTAAATCCAGAACAGTTAGCTAAATCAAAATCATGGACTATTGATAGTTTCATTGAG
AAAAAAGCAATTTTTATCACAATTATTGCATTTTTGATGGGTATCTCCTATGCTTCCGTGTTA
GGTTTCCAAAAATTATATACAACAGAAATTAATTTGATGACAGTAGGAGCTTATTTCTTTATT
GTTTATGCACTTGTCATCACTTTAACCAGACCATCTATGGGAAGATTAATGGACGCTAAGGGA
GATAAGTGGGTGCTTTATCCAAGTTATCTGTTCTTAACTTTGGGACTTGCTTTATTAGGGAGT
GCTATGGGAAGTGTTACCTACCTTCTATCAGGTGCTTTGATTGGTTTTGGTTATGGCACCTTT
ATGTCTTGTGGCCAAGCAGCATCAATCAAAGGTGTTGAGGAACATCGTTTCAATACAGCCATG
TCAACTTACATGATAGGTCTTGATTTAGGGTTAGGTGCTGGACCTTACATTTTGGGACTTGTT
AAAGATGGTTTTCTTGGAGCTGGTGTGCAATCCTTTAGAGAATTATTCTGGATAGCAGCGATT

ATTCCTGTTGTTTGTGGTATTCTATATTTCTTAAAATCATCTAGACAAGTTGAAACTAAAACT
ATATAA

(B)
MEDKLFNKHFIGITILNFIVYMVYYLFTVIIAFIATKELGVSTSQAGLATGIYIVGTLIARLI
FGKQLEVLGRKLVLRGGAIFYLLTTLAYFYMPSIGVMYLVRFLNGFGYGVVSTATNTIVTAYI
PADKRGEGINFYGLSTSLAAAIGPFVGTFMLDNLHINFKMVIVLCSILIAIVVLGAFVFPVKN
ITLNPEQLAKSKSWTIDSFIEKKAIFITIIAFLMGISYASVLGFQKLYTTEINLMTVGAYFFI
VYALVITLTRPSMGRLMDAKGDKWVLYPSYLFLTLGLALLGSAMGSVTYLLSGALIGFGYGTF
MSCGQAASIKGVEEHRFNTAMSTYMIGLDLGLGAGPYILGLVKDGFLGAGVQSFRELFWIAAI
IPVVCGILYFLKSSRQVETKTIZ

ID-19: 927 base pairs
Clone 102
(A)
ATGAAAAAGATTCGATTATCAAAGTTTATTAAAATGATTGTTGTTATTTTGTTTTTAATTAGT
GTAGCAGCTAGTTTTTATTTTTTCCACGTTGCCCAAGTTCGAGATGATAAATCCTTTATTTCA
AATGGTCAACGTAAGCCTGGAAACTCTTTATATGCTTATGATAAATCCTTTGATAAGCTATTA
AAGCAAAAAATAGAAATGACAAACCAAAATATAAAGCAAGTTGCTTGGTATGTTCCTGCTGCT
AAGAAAACTCATAAGACAGTTGTTGTCGTTCATGGTTTTGCGAATAGCAAAGAGAATATGAAG
GCATATGGTTGGCTGTTTCATAAGTTAGGATACAATGTTCTTATGCCTGACAACATTGCACAT
GGTGAAAGTCATGGGCAGTTGATAGGCTATGGCTGGAACGACCGCGAGAACATTATCAAATGG
ACAGAAATGATAGTGGATAAGAATCCATCAAGCCAAATTACTTTATTTGGTGTTTCAATGGGT
GGAGCAACAGTCATGATGGCTAGTGGTGAAAAATTACCTAGTCAGGTTGTTAATATCATTGAA
GATTGTGGTTATTCTAGTGTTTGGGATGAATTAAAATTTCAGGCTAAAGAGATGTATGGTTTA
CCAGCCTTCCCACTCTTATATGAAGTTTCAACAATTTCTAAAATCAGAGCAGGTTTTTCGTAT
GGACAAGCAAGTAGTGTCGAACAATTGAAAAAGAATAATTTACCAGCCCTCTTTATTCATGGT
GATAAGGATAATTTTGTTCCAACAAGTATGGTTTATGACAACTATAAAGCTACAGCAGGTAAG
AAAGAGCTTTATATTGTAAAAGGGGCAAAACATGCGAAATCTTTTGAAACAGAGCCAGAAAAA
TATGAGAAACGTATCTCTAGTTTTTTGAAAAAATATGAAAAATAA

(B)
MKKIRLSKFIKMIVVILFLISVAASFYFFHVAQVRDDKSFISNGQRKPGNSLYAYDKSFDKLL
KQKIEMTNQNIKQVAWYVPAAKKTHKTVVVVHGFANSKENMKAYGWLFHKLGYNVLMPDNIAH
GESHGQLIGYGWNDRENIIKWTEMIVDKNPSSQITLFGVSMGGATVMMASGEKLPSQVVNIIE
DCGYSSVWDELKFQAKEMYGLPAFPLLYEVSTISKIRAGFSYGQASSVEQLKKNNLPALFIHG
DKDNFVPTSMVYDNYKATAGKKELYIVKGAKHAKSFETEPEKYEKRISSFLKKYEKZ

ID-20: 546 base pairs
Clone 120
(A)
TTGAGGAGTAATATGGTAAAGACAGCAGTTTTAATGGCGACATACAATGGCGAAAAATTTATA
TCTGAACAACTTGATTCAATTCGCCAACAGACATTAAAACCAGATTATGTATTATTGAGGGAT

GATTGTTCAACGGATGAAACAGTCAATGTCGTCAATAACTATATCGCAAAACATGAGTTAGAA
GGCTGGAAAATTGTTAAAAACGACAAAAACTTAGGCTGGCGTTTAAATTTTCGTCAATTACTT
ATTGATGTGTTAGCCTATGAGGTTGACTATGTCTTTTTTAGTGATCAAGATGATATTTGGTAT
CTTGATAAAAACGAACGACAGTTTGCCATTATGTCAGATAACCCTCAAATTGAGGTTTTGAGT
GCAGACGTTGATATCAAAACGATGTCTACAGAAGCCAGTGTTCCACATTTTCTAACTTTTTCT
TCTAGTGATAGAATCAGTCAGTATCCTAAAGTATATGATTATCAAACATTCCGTCCCGGATGG
ACCATTGCTATGAAGAGAGATTTTGCGCAAGCTATCGCTTGA

(B)
MRSNMVKTAVLMATYNGEKFISEQLDSIRQQTLKPDYVLLRDDCSTDETVNVVNNYIAKHELE
GWKIVKNDKNLGWRLNFRQLLIDVLAYEVDYVFFSDQDDIWYLDKNERQFAIMSDNPQIEVLS
ADVDIKTMSTEASVPHFLTFSSSDRISQYPKVYDYQTFRPGWTIAMKRDFAQAIAZ

ID-21: 579 base pairs
Clone 143
(A)
ATGATTCATGAGATTCACGATTGTCAATTTATTGAAAAAGGAAGTTACGTTTATTTGAATTAT
ATTAATGCTGAGGGCGAGAGAGTAGTTATTATAATCATAGATTTTGTCCGTAGTGTTAGTCCT
ATTTTATATCGTCTATTTATGATTTTACTTGCACAAGAAGTACCTCACTTGCATGATTACATC
TATAATGCAAGAGATGATCACTACGATACTTGGAAGTTTAAAGAATTAAAGGAGTCAAACCAT
CCAGTCCTTTTGGCATTCTCTGAAAGGTGGCACGATAGTCGCTTGACTTCTAAAAGCCTTGCA
GAATGTTTACAATTAACCGACCTTGATGAAGAAGTGAAATCGACCATCATTCAATTAAGACAG
TTCGAAAAATCAGTCAGAAATCCTTTGGCTCACCTGATTAAACCTTTTGATGAGCAAGAACTA
TATCGTACAACTCAATTTTCTTCTCAAGCATTTTTAGACCAGATTATCTTCTTGGCAAAGGTA
ATTGGTGTTGAGTATGATACTGTTAATTTTCACTACGATACGGTTAACAAGCTTATTATAAAG
ATACTTGAGTAA

(B)
MIHEIHDCQFIEKGSYVYLNYINAEGERVVIIIIDFVRSVSPILYRLFMILLAQEVPHLHDYI
YNARDDHYDTWKFKELKESNHPVLLAFSERWHDSRLTSKSLAECLQLTDLDEEVKSTIIQLRQ
FEKSVRNPLAHLIKPFDEQELYRTTQFSSQAFLDQIIFLAKVIGVEYDTVNFHYDTVNKLIIK
ILEZ

FIGURE 2

ID-4:
Clone 6b
(A)
TTGATGAAGTCTAATCAATGGCAAGTCTTTAAGAGATTAATCTCCTATTTACGCCCTTATAAA
TGGTTTACAGTATTAGCTCTATCTCTCTTATTGTTGACGACTGTTGTTAAAAATATTATTCCT
TTAATTGCTTCACATTTTATTGATCACTATCTGACAAATGTTAATCAAACAGCAGTTCTTATT
TTAGTGGGATATTATTCAATGTATGTCTTGCAGACCTTAATTCAATATTTTGGGAATCTCTTT
TTTGCGCGTGTTTCTTATAGTATTGTTAGAGATATTCGTAGAGATGCTTTTGCTAATATGGAA
AGGCTAGGCATGTCTTATTTTGATAGGACACCGGCAGGATCTATTGTGTCACGTATTACTAAT
GATACTGAAGCAATATCTGATATGTTTTCGGGTATTTTATCAAGTTTTATCTCGGCGATATTT
ATTTTTACAGTTACTCTGTACACTATGTTGATGCTAGACATTAAACTAACAGGACTCGTCGCT
CTTTTGTTACCTGTTATCTTTATATTAGTGAATGTCTATCGGAAAAAATCAGTCACTGTCATT
GCTAAAACGAGAAGTTTACTTAGTGATATCAACAGTAAATTATCAGGAAGTATTGAAGGAATT
CGCATTGTACAGGCTTTTGGTCAAGAAGAGCGCTTGAAGACTGAATTTGAGGAAATTAACAAA
GAGCATGTTGTGTATGCCAATCGTTCTATGGCTCTTGATAGTCTCTTCTTAAGACCGGCGATG
TCTCTTTTAAAACTCCTAGCATATGCTGTTCTTATGTCTTATTTTGGATTTACAGGAGTTAAA
GGAGGTCTTACGGCAGGATTAATGTATGCTTTTATTCAGTACGTTAATCGTCTATTTGACCCT
TTAATTGAAGTAACGCAAAATTTTTCAACCTTACAAACATCAATGGTATCAGCAGGGCGTGTG
TTTGATCTGATTGATGAAACAGGTTTTGAACCAAGCCAAAAAAATACAGAAGCT

(B)
MMKSQWQVFKRLISYLRPYKWFTVLALSLLLLTTVVKNIIPLIASHFIDHYLTNVNQTAVLIL
VGYYSMYVLQTLIQYFGNLFFARVSYSIVRDIRRDAFANMERLGMSYFDRTPAGSIVSRITND
TEAISDMFSGILSSFISAIFIFTVTLYTMLMLDIKLTGLVALLLPVIFILVNVYRKKSVTVIA
KTRSLLSDINSKLSGSIEGIRIVQAFGQEERLKTEFEEINKEHVVYANRSMALDSLFLRPAMS
LLKLLAYAVLMSYFGFTGVKGGLTAGLMYAFIQYVNRLFDPLIEVTQNFSTLQTSMVSAGRVF
DLIDETGFEPSQKNTEA

ID-5: 654- base pairs
Clone 7
(A)
ATGAAAAGAAAAGACTTATTTGGTGATAAACAAACTCAATACACGATTAGAAAGTTAAGTGTT
GGAGTAGCTTCAGTTGCAACAGGGGTATGTATTTTTCTTCATAGTCCACAGGTATTTGCTGAA
GAAGTAAGTGTTTCTCCTGCAACTACAGCGATTGCAAAGTCGAATATTAATCAGGTTGACAAC
CGGCAATCTACTAATTTAAAAGATGACATAAACTCAAACTCTGAGACGGTTGTGACACCCTCA
GATATGCCGGATACCAAGCAATTAGTATCAGATGAAACTGACACTCAAAAAGGAGTGACAGAG
CCGGATAAGGCGACAAGCCTGCTTGAAGAAAATAAAGGTCCTGTTTCAGATAAAAATACCTTA
GATTTAAAAGTGGCACCATCTACATTGCAAAATACTCCCGACAAAACTTCTCAAGCTATAGGT
GCTCCAAGTCCGACCTTGAAAGTTGCTAATCAAGCTCCACAGATTGAAAATGGTTACTTTAGG
TTACATCTTAAAGAATTGCCTCAAGGTCATCCTGTAGAAAGCACTGGGCTTTGGATATGGGGA

GATGTTGATCAACCGTCTAGTAATTGGCCAAATGGTGCTATCCCTATGACTAATGCTAAGAAA
GATGATTACGGTTATTATGCTTGA

(B)
MKRKDLFGDKQTQYTIRKLSVGVASVATGVCIFLHSPQVFAEEVSVSPATTAIAKSNINQVDN
RQSTNLKDDINSNSETVVTPSDMPDTKQLVSDETDTQKGVTEPDKATSLLEENKGPVSDKNTL
DLKVAPSTLQNTPDKTSQAIGAPSPTLKVANQAPQIENGYFRLHLKELPQGHPVESTGLWIWG
DVDQPSSNWPNGAIPMTNAKKDDYGYYAZ

ID-7: 528- base pairs
Clone 15
(A)
TTGTTCAATAAAATAGTTTTAGAACTTGGAAATCAGGAAAGCTTTTGGCTTTATATGGGAGTG
CTAGGATCAACTATTATTTTAGGATCAAGTCCTGTATCTGCTATGGATAGTGTTGGAAATCAA
AGTCAAGGTAATGTTTTAGAGCGTCGCCAACGTGATGCGGAAAACAAAAGTCAGGGTAATGTT
TTAGAGCGTCGCCAACGTGATGCGGAAAACAAGAGCCAAGGCAATGTTTTAGAGCGTCGTCAA
CGCGATGTTGAGAATAAGAGCCAAGGCAATGTTTTAGAGCGTCGTCAACGTGATGCGGAAAAC
AAAAGTCAGGGCAATGTTCTAGAGCGCCGCCAACGTGATGCGGATAACAAGAGCCAAGTAGGT
CAACTTATAGGGAAAAATCCACTTTTTTCAAAGCCAACTGTATCTAGAGAAAATAATCACTCT
AGTCAAGGTGACTCTAACAAACAGTCATTCTCTAAAAAAGTATCTCAGGTTACTAATGTAGCT
AATAGACCGATGTTAATCCAT

(B)
MFNKIVLELGNQESFWLYMGVLGSTIILGSSPVSAMDSVGNQSQGNVLERRQRDAENKSQGNV
LERRQRDAENKSQGNVLERRQRDVENKSQGNVLERRQRDAENKSQGNVLERRQRDADNKSQVG
QLIGKNPLFSKPTVSRENNHSSQGDSNKQSFSKKVSQVTNVANRPMLIH

ID-11: 942 base pairs
Clone 23
(A)
ATGACTTATCAAAAAACAGTTGTTTTGGCTGGTGATTATTCCTACATTAGACAAATTGAAACC
ACATTAAAATCTCTCTGTGTCTATCATGAGAATCTCTCAATTTTTATTTTTAATCAAGATATT
CCTCAAGAATGGTTTTTAGCTATGAAAGATAGGGTTGGACAAACTGGAAATCAAATTCAGGAT
GTAAAGCTCTTCCATGATCACTTATCCCCAAAATGGGAAAATAAAAAGCTTAATCATATTAAT
TATATGACCTATGCTCGTTATTTCATACCTCAGTACATCTCAGCTGATACAGTTTTATATCTT
GACTCTGACTTAGTTGTTACTACTAATTTAGATAACCTCTTTCAAATTTCACTAGACAATGCA
TATTTAGCTGCAGTTCCAGCTCTTTTTGGGCTTGGATATGGGTTTAATGCTGGAGTAATGGTA
ATTAACAACCAACGTTGGCGACAAGAAAATATGACTATTAAATTAATTGAAAAAAATCAAAAG
GAAATTGAGAATGCCAACGAAGGGGATCAAACAATTCTTAATCGCATGTTTGAAAATCAGGTA
ATTTATTTAGATGATACCTACAATTTTCAAATTGGTTTTGATATGGGAGCTGCTATCGATGGG
CATAAATTTATTTTTGACATCCCAATTACCCCACTCCCAAAAATTATTCACTACATTTCGGGA
ATCAAACCTTGGCAAACATTATCAAATATGAGACTCCGTGAGGTATGGTGGCACTATAATTTA
CTTGAATGGTCAAGTATCATATCTAGTAAAAAAGTATTTGGTTTAGACCACCCAATTAAAACA

CAAAATTATCGTCTCAATTTCCTTATTGCTACAACTTCTGATTGTATACCATCTATCTCAGAA
TTAGTCACTGCCCTTCCAGATTGTCTATTTCACATTGCATGCACCAACAGTTATGTCTGA

(B)
MTYQKTVVLAGDYSYIRQIETTLKSLCVYHENLSIFIFNQDIPQEWFLAMKDRVGQTGNQIQD
VKLFHDHLSPKWENKKLNHINYMTYARYFIPQYISADTVLYLDSDLVVTTNLDNLFQISLDNA
YLAAVPALFGLGYGFNAGVMVINNQRWRQENMTIKLIEKNQKEIENANEGDQTILNRMFENQV
IYLDDTYNFQIGFDMGAAIDGHKFIFDIPITPLPKIIHYISGIKPWQTLSNMRLREVWWHYNL
LEWSSIISSKKVFGLDHPIKTQNYRLNFLIATTSDCIPSISELVTALPDCLFHIACTNSYVZ

ID-12: 1146 base pairs
Clone 27
(A)
GTGAAGAAAACATATTGTTATATCGGCTCAGTTGCTGCTATTTTACTAGCTACTCATATTGGA
AGTTACCAGCTTGGTAAGCATCATATGGGTCTAGCAACAAAGGACAATCAGATTGCCTATATT
GATGATAGCAAAGGTAAGGTAAAAGCCCCTAAAACAAACAAAACGATGGATCAAATCAGTGCT
GAAGAAGGCATCTCTGCTGAACAGATCGTAGTCAAAATTACTGACCAAGGTTATGTTACCTCA
CACGGTGACCATTATCATTTTTACAATGGGAAAGTTCCTTATGATGCGATTATTAGTGAAGAG
TTGTTGATGACGGATCCTAATTACCATTTTAAACAATCAGACGTTATCAATGAAATCTTAGAC
GGTTACGTTATTAAAGTCAATGGCAACTATTATGTTTACCTCAAGCCAGGTAGTAAGCGCAAA
AACATTCGAACCAAACAACAAATTGCTGAGCAAGTAGCCAAAGGAACTAAAGAAGCTAAAGAA
AAAGGTTTAGCTCAAGTGGCCCATCTCAGTAAAGAAGAAGTTGCGGCAGTCAATGAAGCAAAA
AGACAAGGACGCTATACTACAGACGATGGCTATATTTTTAGTCCGACAGATATCATTGATGAT
TTAGGAGATGCTTATTTAGTACCTCATGGTAATCACTATCATTATATTCCTAAAAAAGATTTG
TCTCCAAGTGAGCTAGCTGCTGCACAAGCCTACTGGAGTCAAAAACAAGGTCGAGGTGCTAGA
CCGTCTGATTACCGCCCGACACCAGCCCCAGGTCGTAGGAAAGCCCCACTTCCTGATGTGACG
CCTAACCCTGGACAAGGTCATCAGCCAGATAACGGTGGTTATCATCCAGCGCCTCCTAGGCCA
AATGATGCGTCACAAAACAAACACCAAAGAGATGAGTTTAAAGGAAAAACCTTTAAGGAACTT
TTAGATCAACTACACCGTCTTGATTTGAAATACCGTCATGTGGAAGAAGATGGGTTGATTTTT
GAACCGACTCAAGTGATCAAATCAAACGCTTTTGGGTATGTGGTGCCTCATGGAGATCATTAT
CATATTATCCCAAGAAGTCAGTTATCACCTCTTGAAATGGAATTAGCAGATCGATACTTAACC
CGGCCAAACTGA

(B)
MKKTYCYIGSVAAILLATHIGSYQLGKHHMGLATKDNQIAYIDDSKGKVKAPKTNKTMDQISA
EEGISAEQIVVKITDQGYVTSHGDHYHFYNGKVPYDAIISEELLMTDPNYHFKQSDVINEILD
GYVIKVNGNYYVYLKPGSKRKNIRTKQQIAEQVAKGTKEAKEKGLAQVAHLSKEEVAAVNEAK
RQGRYTTDDGYIFSPTDIIDDLGDAYLVPHGNHYHYIPKKDLSPSELAAAQAYWSQKQGRGAR
PSDYRPTPAPGRRKAPLPDVTPNPGQGHQPDNGGYHPAPPRPNDASQNKHQRDEFKGKTFKEL
LDQLHRLDLKYRHVEEDGLIFEPTQVIKSNAFGYVVPHGDHYHIIPRSQLSPLEMELADRYLT
RPNZ

ID-18: 414 base pairs
Clone 47
(A)
ATGATACTTGGAGGCTGTCAAATGAATAGTGAACCTAAAAGTCAGTCAAACGAAGTAAAAAAT
AGCAAGCAATCAGAAGTGAAGAAAGATAAAAAAATGACAAAAAAAGAACAATTAGCCTATCTC
AAAGAGCATGAGCAAGAAATCATAGATTATGTAAAATTACATAACAACCAAATTGAGTCCGTT
CAATTCGATTGGTCAAGTGTAAAAGTAGAACAAAGCGGGAATGGAACTCCACAAGGGGGTGAT
TATAATCTTTCACTGAGAGGAAAGTTTAATCATCTACAAAATTCAAAATTAATAGTTGATTTT
TATTTAGCTCATAAAAATGATATCCCAAATATCAAATCAATGGGAATGCTAAATAAGCCATAT
ATACATAAAAATGGTATTTGGCACATTTATGAATAG

(B)
MILGGCQMNSEPKSQSNEVKNSKQSEVKKDKKMTKKEQLAYLKEHEQEIIDYVKLHNNQIESV
QFDWSSVKVEQSGNGTPQGGDYNLSLRGKFNHLQNSKLIVDFYLAHKNDIPNIKSMGMLNKPY
IHKNGIWHIYEZ

ID-22: 477 base pairs
Clone 1
(A)
ATGGTAAAAGTTTCAAATTTAGGGTATCCACGTCTTGGTGAACAGCGCGAATGGAAGCAAGCG
ATCGAAGCTTTCTGGGCAGGGAATCTTGAACAAAAAGATTTAGAAAAACAACTAAAACAATTA
CGTATCAATCATTTAAAGAAACAAAAAGAGGCAGGTATTGACCTTATTCCAGTGGGGGATTTT
TCTTGTTATGATCATGTTTTGGATTTGTCATTTCAATTCAATGTAATCCCAAAGCGTTTCGAT
GAGTATGAGAGGAATTTAGACCTTTATTTTGCTATTGCAAGAGGTGACAAAGATAATGTCGCA
TCATCTATGAAAAAGTGGTTTAATACCAACTACCACTACATAGTCCCAGAATGGGAGGTTGAG
ACTAAACCTCACTTGCAGAATAATTACTTACTTGATCTTTATCTAGAAGCTAGGGAAGTAGTT
GGTGATAAAGCAAAGCCGGTTATC

(B)
MEEIMVKVSNLGYPRLGEQREWKQAIEAFWAGNLEQKDLEKQLKQLRINHLKKQKEAGIDLIP
VGDFSCYDHVLDLSFQFNVIPKRFDEYERNLDLYFAIARGDKDNVASSMKKWFNTNYHYIVPE
WEVETKPHLQNNYLLDLYLEAREVVGDKAKPVI

ID-23: 124 base pairs
Clone 2
(A)
ATGGTGTTACTTTTATTGCTAATGGTAGCCAAGTCAAGTTTGATGGTTACATGGCTGTTTATA
ACGATACTGACAAAAATAAAATGTTACCAGATATGGAGGAAGGAGAAAGTTATCAAGTTAA

(B)
MVLLLLLMVAKSSLMVTWLFITILTKIKCYQIWRKEKVIKL

ID-24: 158 base pairs
Clone 14
(A)
ATGAACAAAAAAATTTCCGGGATCGGCTTGGCTTCGATTGCAGTACTTAGTTTAGCTGCATGT
GGACATCGTGGTGCTTCTAAATCTGGTGGTAAATCAGATAGCTTGAAGGTTGCAATGGTAACA
GATACCGGTGGTGTTGATGATAAATCATTTAA

(B)
MNKKISGIGLASIAVLSLAACGHRGASKSGGKSDSLKVAMVTDTGGVDDKSF

ID-25: 240 base pairs
Clone 20
(A)
GTGAGTTTTTATATGTTACATTCTAAAAAAATACATTCCTTATCGCTTATTGCCGTTCTCTCT
TTAGCAACATATACGAGTTTACAACCAAATCATGTAGCGGCTGAACAATCACAAAAAACATCA
ACTGTTCATATGAGTCAAAAAACTATTGAACATAAGTTAAAAGTTGCAGATAAAGAAGCTGCT
CCTCTCTACGCTAAAATCGACCATATCCAACGACATATTGAAGTCAAAAAAGCAAGAGATTTA
A

(B)
MSFYMLHSKKIHSLSLIAVLSLATYTSLQPNHVAAEQSQKTSTVHMSQKTIEHKLKVADKEAA
PLYAKIDHIQRHIEVKKARDL

ID-26: 465 base pairs
Clone 25
(A)
CTGAATTCCCAAAAACGCTACAATCAAACTTGGTATCCTACTTATGGTTTTTCTGATACTTAT
GCATTCATGGTTACTAAAGAGTTTGCCAGACAGAATAAAATCACCAAGATCTCTGATCTCAAA
AAGTTATCAACAACTATGAAGGCAGGGGTTGATAGTTCATGGATGAATCGCGAGGGAGATGGA
TACACTGATTTCGCTAAAACATACGGTTTTGAATTTTCACATATTTACCCTATGCAAATTGGC
TTAGTCTATGATGCGGTTGAAAGTAACAAAATGCAATCTGTATTAGGCTACTCCACTGACGGT
CGTATTTCGAGCTATGATTTAGAAATTTTAAGGGATGATAAAAAATTCTTTCCTCCTTATGAA
GCCTCTATGGTTGTCAACAATTCTATCATCAAAAAAGATCCTAAACTAAAAAAATTACTCCAT
CGACTCGATGGTAAAATCAATTTA

(B)
MNSQKRYNQTWYPTYGFSDTYAFMVTKEFARQNKITKISDLKKLSTTMKAGVDSSWMNREGDG
YTDFAKTYGFEFSHIYPMQIGLVYDAVESNKMQSVLGYSTDGRISSYDLEILRDDKKFFPPYE
ASMVVNNSIIKKDPKLKKLLHRLDGKINL

ID-28: 125 base pairs
Clone 34
(A)
ATGACAAAAAAACTTATTATTGCTATATTAGCACTATGCACTATCTTAACCACTTCTCAAGCT
GTTTTAGCTAAAGAAAAATCACAAACTGTTACCATAAAAAACAACTATTCGGTCTATATTAA

(B)
MTKKLIIAILALCTILTTSQAVLAKEKSQTVTIKNNYSVYI

ID-29: 188 base pairs
Clone 37
(A)
ATGAAAAAATTACTTTCCCTAACATGTCTAATCATGATGTCTTTATGTTTAGTGGCATGTACT
AAGCAAGCAATGTCGTCTAAGCAAGCAATGTCGTCTAAGCAAATTAAAGATAAGAATAGTAAA
GAAAAGGTGATTACTGTTGCAACTTACAGCAAACCTACATCTACCTTTTTAGATTTGATTAA

(B)
MKKLLSLTCLIMMSLCLVACTKQAMSSKQAMSSKQIKDKNSKEKVITVATYSKPTSTFLDLI

ID-30: 711 base pairs
Clone 38
(A)
CTGTTGGCTAAGGAAACCACTATGTCTGTCCTTTGGTATCAAAATTCTGCAGAAGCCAAGGCT
TTATATTTACAAGGTTATAATGTTGCTAAAATGAAGTTAGATGATTGGTTACAAAAGCCCAGT
GAAAAACCATATTCAATTATCTTAGATTTAGATGAAACAGTTTTAGATAATAGCCCATATCAA
GCAAAGAATATTAAAGATGGCTCTAGTTTCACGCCAGAGAGTTGGGATAAATGGGTGCAAAAG
AAATCAGCTAAGGCTGTTGCGGGTGCCAAAGAATTTTTGAAGTATGCTAATGAAAAGGGAATA
AAAATTTATTATGTCTCAGATCGTACAGATGCTCAAGTTGATGCGACTAAAGAAAATTTAGAG
AAGGAAGGTATACCTGTTCAAGGGAAAGACCACTTGCTTTTCCTTAAAAAAGGAATGAAATCT
AAAGAGAGTCGCCGTCAGGCAGTTCAAAAAGATACCAATTTAATTATGCTTTTTGGAGATAAT
TTAGTTGATTTTGCTGATTTTTCTAAATCATCTAGTACAGATAGAGAACAACTACTAACTAAA
CTTCAAAGTGAGTTTGGTAGTAAATTTATTGTTTTCCCAAATCCTATGTACGGTTCTTGGGAA
AGTGCTATTTATCAAGGAAAACATCTGGATGTTCAAAAACAATTGAAAGAACGACAAAAAATG
TTGCATTCGTATGATTAA

(B)
MLAKETTMSVLWYQNSAEAKALYLQGYNVAKMKLDDWLQKPSEKPYSIILDLDETVLDNSPYQ
AKNIKDGSSFTPESWDKWVQKKSAKAVAGAKEFLKYANEKGIKIYYVSDRTDAQVDATKENLE
KEGIPVQGKDHLLFLKKGMKSKESRRQAVQKDTNLIMLFGDNLVDFADFSKSSSTDREQLLTK
LQSEFGSKFIVFPNPMYGSWESAIYQGKHLDVQKQLKERQKMLHSYDZ

ID-31: 128 base pairs
Clone 41
(A)
ATGGATAATAAAGGTAATAACGCCAATGTGATTGATGCAATCGCTGAGGGTGCAAGCACAGGT
GCACAAATGGCTTTCTCAATTGGTGCTAGTTTGATTGCCTTTGTTGGTTTAGTTTCTTTGATT
AA

(B)
MDNKGNNANVIDAIAEGASTGAQMAFSIGASLIAFVGLVSLI

ID-32: 116 base pairs
Clone 42
(A)
ATGAAAAAGAAAAACAAATCCTCTAACATTGCTATAATTGCAATCTTTTTTGCTATTATGCTT
GTCATTCATTTTTTGTCATCATTTATTTTTAGTTTTTGGTTAGTCCCTATTAA

(B)
MKKKNKSSNIAIIAIFFAIMLVIHFLSSFIFSFWLVPI

ID-33: 251 base pairs
Clone 43
(A)
TTGAATATGACATTACAAGACGAAATCAAAAAACGCCGTACTTTTGCCATCATCTCTCACCCG
GATGCTGGTAAGACGACTATTACTGAGCAATTATTATATTTTGGTGGTGAAATTAGAGAAGCA
GGGACAGTAAAAGGGAAAAAATCAGGTACTTTTGCAAAGTCCGACTGGATGGATATTGAAAAG
CAACGGGGTATCTCTGTTACTTCATCTGTTATGCAATTTGATTACGCGGGTAAACGTGTTAA

(B)
MNMTLQDEIKKRRTFAIISHPDAGKTTITEQLLYFGGEIREAGTVKGKKSGTFAKSDWMDIEK
QRGISVTSSVMQFDYAGKRV

ID-34: 296 base pairs
Clone 44
(A)
ATGGCAGATAAAAACAGAACATTTAAACTTGTAGGTGCAGGATCTTCTAGCACACAAGAAAAA
ATTGAAAAGCCTGCTCTTTCGTTTATGCAAGATGCGTGGCGTCGCTTGAAAAAAAACAAATTA
GCAGTAGTTTCACTCTATTTATTAGCTCTTTTACTTACTTTTTCGTTAGCCTCAAATTTATTT
GTAACTCAGAAGGATGCTAATGGGTTTGATTCGAAAAAAGTAACGACATATCGCAACTTACCA
CCTAAATTGAGTTCAAACCTTCCTTTTTGGAATGGTAGCATTAA

(B)
MADKNRTFKLVGAGSSSTQEKIEKPALSFMQDAWRRLKKNKLAVVSLYLLALLLTFSLASNLF
VTQKDANGFDSKKVTTYRNLPPKLSSNLPFWNGSI

ID-35: 154 base pairs
Clone 46
(A)
ATGAAAAGAAAACAGTTTATAAAATTAGGAATTGCAACCTTACTAACGGTTATTTCGCTTTAC
ACACCAATAAACCTAGCTACAAATCATACCACAGAAAATATTGTTACTGCTCAAGAGTATAAA
ACAAAGAGAATGGTACTTTACCTTTTAA

(B)
MKRKQFIKLGIATLLTVISLYTPINLATNHTTENIVTAQEYKTKENILFLL

ID-36: 143 base pairs
Clone 50
(A)
ATGTTTTATAATCCTTTACTTTTTATTGTACTAATTACAATTGCTGTATTTTTCTTAGCTAAG
AAAAAATGGCAATTACCGACATTTACTTTCATTGGTTTGCTATTTATCTATAACCAAGGGCTG
TGGGAACAGTTGATTAAT

(B)
MFYNPLLFIVLITIAVFFLAKKKWQLPTFTFIGLLFIYNQGLWEQLIN

ID-37: 338 base pairs
Clone 51/52
(A)
GTGGTGCAAATAATGAAAAAACATATAAAAAGTATCATACCAATAGTTCTTATTGGTATGATA
CTAGGAGGCTGTCAAATGAATAGTGAACATAAAAGTCAGTATAATGAAACAAAAAGTAGCAAG
CAATCAGAAGTGAAGAAAGATAAAAAAATGACAAAAAAAGAACAATTAGCTTATCTCAAAGAG
CATGAACAAGAAATAATTGATTTTGTAAAATCTCAGAATAAAAAGATAGAATCTGTACAAATT
GATTGGAATGATGTTCGATGGAGTAAAGGGGGAAATGGTACACCTCAAGGAGGAGGAGAGGGG
ATTTTACTTTTTGGGGAGATTAA

(B)
MVQIMKKHIKSIIPIVLIGMILGGCQMNSEHKSQYNETKSSKQSEVKKDKKMTKKEQLAYLKE
HEQEIIDFVKSQNKKIESVQIDWNDVRWSKGGNGTPQGGGEGILLFGEI

ID-38: 374 base pairs
Clone 53
(A)
ATGGAATTTTTGGCTTATAATGCTTTCACAGCAATCGGTGTTTCTATTCCGCACGGTAATCAT
TTCCACTTTATTCACTATAAGGATATGTCTCCATTAGAGTTAGAAGCAACAAGGATGGTGGCA
GAGCATAGAGGACATCATATTGATGCATTAGGGAAAAAAGATTCTACAGAGAAACCAAAGCAT
ATTTCTCATGAACCTAATAAGGAACCTCACACAGAGGAAGAACACCATGCAGTAACACCGAAA
GACCAACGTAAAGGCAAACCAAATAGCCAGATTGTCTACAGTGCTCAAGAAATTGAAGAGGCA
AAAAAAGCTGGTAAATACACAACATCTGATGGTTACATTTTTGATGCTAAAGATATTAA

(B)
MEFLAYNAFTAIGVSIPHGNHFHFIHYKDMSPLELEATRMVAEHRGHHIDALGKKDSTEKPKH
ISHEPNKEPHTEEEHHAVTPKDQRKGKPNSQIVYSAQEIEEAKKAGKYTTSDGYIFDAKDI

ID-39: 182 base pairs
Clone 56
(A)
ATGAGGAAACGTTTTTCCTTGCTAAATTTTATTGTTGTTACTTTTATTTTCTTTTTCTTTATT
CTTTTTCCGCTTTTTAAGGCCAAAGATTGTCAGGTTGTTTATGCAAGTTTTCAAGGAGATCAT
TGGGACATTTGTAACGCATTTGATTTTCCGTATTTACATCGCTTTGATCTCATTAA

(B)
MRKRFSLLNFIVVTFIFFFFILFPLFKAKDCQVVYASFQGDHWDICNAFDFPYLHRFDLI

ID-40: 948 base pairs
Clone 57
(A)
TTAAATGCTGTCCAATCTGGGCAAGCTGACGGTGTTATTGCAGGAGCCACAATCACAGAAGCA
CGCCAAAAAATCTTTGATTTTTCTGATCCTTATTACACATCTAGCGTTATCTTAGCGGTTAAA
AAAGGAAGCAATGTCAAATCATACCAAGATTTAAAAGGAAAAACAGTTGGTGCTAAAAATGGT
ACTGCCTCATATACTTGGTTATCAGACCACGCAGATAAGTACAACTATCATGTTAAAGCATTT
GATGAAGCATCTACAATGTATGATAGTATGAACTCAGGTTCAATTGATGCTCTAATGGATGAC
GAAGCCGTTCTTGCTTACGCTATTAATCAAGGTCGTAAATTTGAAACACCTATCAAAGGTGAA
AAATCAGGCGATATCGGATTTGCAGTGAAAAAAGGGGCAAATCCAGAATTAATTAAAATGTTT
AACAACGGTCTTGCTTCACTCAAAAAATCGGGTGAGTACGATAAACTTGTTAAAAAATACCTT
TCCACAGCCAGCACTTCTTCAAACGATAAAGCTGCTAAACCTGTAGATGAATCAACTATTTTA
GGGTTAATTTCTAATAACTACAAACAATTGCTATCTGGTATTGGAACTACTTTAAGTTTAACT
CTTATCTCGTTTGCGATTGCTATGGTTATTGGTATTATCTTTGGTATGATGAGCGTATCACCA
AGTAATACTCTCCGCACAATTTCAATGATTTTTGTTGATATTGTCCGTGGTATTCCACTCATG
ATTGTGGCCGCTTTTATTTTCTGGGGTATTCCTAATTTAATCGAAAGCATCACAGGTCACCAA
AGTCCAATTAATGACTTCGTTGCTGCTACTATCGCTCTTTCTTTAAATGGGTGGTGCCGTACA
TTGCTGAAATTGTACGTGGTGGTATTGAAGCTGTTCCTTCTGGTCAAATGGAAGCAAGTCGCA
GCT

(B)
LNAVQSGQADGVIAGATITEARQKIFDFSDPYYTSSVILAVKKGSNVKSYQDLKGKTVGAKNG
TASYTWLSDHADKYNYHVKAFDEASTMYDSMNSGSIDALMDDEAVLAYAINQGRKFETPIKGE
KSGDIGFAVKKGANPELIKMFNNGLASLKKSGEYDKLVKKYLSTASTSSNDKAAKPVDESTIL
GLISNNYKQLLSGIGTTLSLTLISFAIAMVIGIIFGMMSVSPSNTLRTISMIFVDIVRGIPLM
IVAAFIFWGIPNLIESITGHQSPINDFVAATIALSLNGWCRTLLKLYVVVLKLFLLVKWKQVA
A

ID-41: 149 base pairs
Clone 58
(A)
TTGGAAGGTTTACTTATTGCATTGATTCCCATGTTTGCGTGGGGAAGTATTGGATTTGTTAGT
AATAAAATTGGAGGGCGTCCAAATCAACAAACATTTGGAATGACTTTAGGAGCATTGCTATTT
GCGATTATCGTATGTTTATTTAA

(B)
MEGLLIALIPMFAWGSIGFVSNKIGGRPNQQTFGMTLGALLFAIIVCLF

ID-42: 963 base pairs
Clone 70
(A)
ATGAATACTATTTATAATACATTGAGAACAGATAAAGGTTATAAAGTTTATGAGGGGTATTTA
TATGAAATTACTGGTGAAGAATGTGAAGAAGCCTTAGACCTTGTGATTCCTAAGAATATTGTA
TTTGCAGATACAGATACTTGTGGCTACACTTTTTTACTCAATGAAGATGGAACAGTTTATGAT
GATGTGACTTTCTACAAATTTGATGATAAATATTGGTTGGCTAGTCATAAAGCTTTGGATTCT
TATTTAGACAACATCAATTTTGACTATACCGTAACAGATATTTCTGACGAGTATAAAATGCTG
CAAATTGAAGGAAGATATTCGGGAGAAATTGCTCAGTCATTTTATGAATATGATATTTCAACA
CTTAATTTTCGTACTCTTCGCATAGAGATGGACTTCATCAAAGGTGAGGAAAGGTTATCTTGG
CGTAGATTTGGTTTTTCTGGAGAATTTGGCTATCAATTTTTCCTACCATCTTCTATTTTTGCT
ACTTTTGTTTCGGATGTCTGTGAAGGTATAGCAGAGTGTGGGGATGAACTTGATAGATATTTA
AGGTTTGAAGTGGGACAACCCATTACTGATATTTATCAACAAGAAGAATATTCTTTATATGAA
ATAGGTTATTCTTGGAATCTAGATTTCACAAAGGAAGAATTTAGAGGTCGCGATAGCTTGTTA
GAGCACATCAGATCAGCAACAGTTAAAAGTGTTGGATTCTCAACGAAGGAAAAACTCGCTTCA
GGAACACCAGTGCTATTTGATGACCAAATTGTTGGAAAGATTTTTTGGATAGCAGACGAGAAA
GACTCTTCGGAAAATTACCTAGGTTTGATGATTGTTAACCAAACATATGCTCATTCAGGAGTT
ACTTTTGTAACAGAAGATGGCCAAATTTTGAAAACACAATCAAGCCCTTATTGTATCCCAGAA
AGTTGGAACAAAGAATGA

(B)
MNTIYNTLRTDKGYKVYEGYLYEITGEECEEALDLVIPKNIVFADTDTCGYTFLLNEDGTVYD
DVTFYKFDDKYWLASHKALDSYLDNINFDYTVTDISDEYKMLQIEGRYSGEIAQSFYEYDIST
LNFRTLRIEMDFIKGEERLSWRRFGFSGEFGYQFFLPSSIFATFVSDVCEGIAECGDELDRYL

RFEVGQPITDIYQQEEYSLYEIGYSWNLDFTKEEFRGRDSLLEHIRSATVKSVGFSTKEKLAS
GTPVLFDDQIVGKIFWIADEKDSSENYLGLMIVNQTYAHSGVTFVTEDGQILKTQSSPYCIPE
SWNKEZ

ID-43: 331 base pairs
Clone 78
(A)
ATGGAGTTAGTAATTAGAGATATTCGTAAGCGGTTTCAGGAAACAGAGGTCTTGAGAGGAGCA
AGTTACCGATTTTATTCAGGTAAAATAACAGGGGTCTTAGGTAGGAATGGTGCTGGGAAAACA
ACTTTATTTATAACTGGTGAGACTGGTGCAGGGAAATCTATCATTATTGATGCTATGAATATG
ATGTTAGGAGCCCGTGCTAGTGTTGAAGTGATTCGCCATGGTGCTAACAAAGCAGAAATTGAA
GGATTTTTCTCTATTGAAAAAAATCAATCATTAGTCCAATTATTGGAAGAAAATGGCATTGAA
TTAGCAGATGAATTAA

(B)
MELVIRDIRKRFQETEVLRGASYRFYSGKITGVLGRNGAGKTTLFITGETGAGKSIIIDAMNM
MLGARASVEVIRHGANKAEIEGFFSIEKNQSLVQLLEENGIELADEL

ID-44: 755 base pairs
Clone 80
(A)
ATGAGATATACAAATGGAAATTTTGAAGCCTTTGCAAGACCTCGAAAACCTGAAGGTGTGGAT
AAAAAATCCGCTTTTATTGTTGGTTCTGGTTTAGCAGGATTAGCTGCCGCTGTCTTTTTAATA
CGTGACGGTCAAATGGATGGTCAACGTATTCATATTTTTGAAGAACTACCTCTTTCTGGAGGA
TCACTTGACGGTGTCCAACGACCTGGATATCGGTTTGGTAACGCGTGGTGGTCGTGAAATGGA
AAATCACTTCGAATGTATGTGGGATATGTACCGTTCCATCCCCTCTCTCGAAGTTCCAGATGC
TTCTTATCTAGATGAATTTTATTGGCTTGACAAGGATGATCCCAATTCATCTAACTGTCGCCT
CATTCATAAACAGGGGAATCGCTTAGAATCTGATGGTGATTTTACACTCGGAACACATTCCAA
AGAGTTAGTTAAGCTAGTCATGGAGACTGAAGAGTCTTTAGGTGCTAAGACGATTGAAGAAGT
TTTTTCAAAAGAATTTTTTGAAAGTAATTTTTGGACTTATTGGGCTACTATGTTTGCCTTTGA
GAAATGGCATTCAGCGATTGAAATGCGTCGATATGCTATGCGCTTTATCCATCATATTTGGTG
GTCTGCCTGATTTCACTTCATTAAAATTTAATAAATATAATCAATATGATTCTATGGTGAAAC
CAATCATCAGTTATTTAGAGTCTCACAATGTAAATGTTCAATTTGATAGCAAGGTAACTAAT

(B)
MRYTNGNFEAFARPRKPEGVDKKSAFIVGSGLAGLAAAVFLIRDGQMDGQRIHIFEELPLSGG
SLDGVQRPGYRFGNAWWSZNGKSLRMYVGYVPFHPLSRSSRCFLSRZILLAZQGZSQFIZLSP
HSZTGESLRIZWZFYTRNTFQRVSZASHGDZRVFRCZDDZRSFFKRIFZKZFLDLLGYYVCLZ
EMAFSDZNASICYALYPSYLVVCLISLHZNLINIINMILWZNQSSVIZSLTMZMFNLIARZL

ID-45: 426 base pairs
Clone 81
(A)
TTGTTGGCTTCTTTATTTATCGTCCGTTTGTCAAAATCGCTTTCGCTAAGGAGGAGCAATATG
AAAAAATTACTTAGATGGCTTCCTCCTGTACTTTTCATTATTATCCTTATAGGAATGACTATC
TTAGGTAAGTCCTATATCAATAAAGTAACAGCTCACAAAATAAAACTCTATAACTCTCGAATG
ACTCCTACTATTTTAATTCCAGGATCCAGTGCTACTCAAGAACGATTTAACAGCATGTTAGCA
CAGCTCAACCAAATGGGAGAAAAACATAGCGTTTTAAAGTTAACTGTCAAAAAAGACAATAGC
ATTATCTACAATGGACAAATTAGCGGCAATGGCCACAAACCCTACCTTGGCATTGGATTTGGA
AATTATGGAGATGGTATTAGAACCATCAAAAACCAACCAAATGGCTAC

(B)
MLASLFIVRLSKSLSLRRSNMKKLLRWLPPVLFIIILIGMTILGKSYINKVTAHKIKLYNSRM
TPTILIPGSSATQERFNSMLAQLNQMGEKHSVLKLTVKKDNSIIYNGQISGNGHKPYLGIGFG
NYGDGIRTIKNQPNGY

ID-46: 401 base pairs
Clone 83
(A)
TTGAAATTAGGTATTACAACATTCGGAGAGACAACAATCCTTGAAGAAACAAACCAAAGCTAT
TCACATCCTGAGAGGATTCGCCAATTAGTTGCTGAGATTGAACTAGCTGATCAAGTTGGTTTA
GATGTATATGGTATTGGAGAGCACCATCGTGAAGATTTTGCGGTCTCTGCACCCGAAATTATC
CTAGCAGCAGGAGCGGTTAGAACTAATAATATCCGTTTATCTAGTGCAGTAACGATTCTCTCT
TCCAATGATCCTATTCGCGTCTATCAGCAATTTTCAACGATTGACGCACTTTCAAATGGTAGA
GCAGAAATTATGGCAGGGCGTGGTTCCTTTATTGAGTCTTTTCCATTGTTTGGATACGATTTA
GCGGATTATGATGATTTATTTAA

(B)
MKLGITTFGETTILEETNQSYSHPERIRQLVAEIELADQVGLDVYGIGEHHREDFAVSAPEII
LAAGAVRTNNIRLSSAVTILSSNDPIRVYQQFSTIDALSNGRAEIMAGRGSFIESFPLFGYDL
ADYDDLF

ID-47: 130 base pairs
Clone 86
(A)
ATGATAGAGTGGATTCAAACACATTTACCAAATGTATATCAAATGGGTTGGGAAGGTGCTTAC
GGCTGGCAGACAGCTATTGTACAAACCCTTTATATGACTTTTTGGTCGTTCCTTATTGGAGGT
TTAA

(B)
MIEWIQTHLPNVYQMGWEGAYGWQTAIVQTLYMTFWSFLIGGL

ID-49: 115 base pairs
Clone 96
(A)
TTGGCAGTTAGTTTTCATGAAGTATTTGGTTGGGATTCTGCTTTTTTTATTATGATTATCAAT
ATTCCATTGCTCCTTCTTTGCTACTTTGGCTTAGGTAAACAAACCTTTTTAA

(B)
MAVSFHEVFGWDSAFFIMIINIPLLLLCYFGLGKQTFL

ID-50: 154 base pairs
Clone 99
(A)
ATGAAAGAAAAACAGTCGAAAAGGCTTATTTATATACTACTGATTGTTCCCATTATCTTTATA
AGTGTTTTTACATACAGTATTAGCCAGCCTTCTAAACTACTTCCACCAAAAGAATTAGTTATT
CTAAGTCCAAATAGTCAAGCCATTTTAA

(B)
MKEKQSKRLIYILLIVPIIFISVFTYSISQPSKLLPPKELVILSPNSQAIL

ID-51: 368 base pairs
Clone 103
(A)
CCTCCTATCAAATGATGACAAACGTGAGAGGTACATGGAACAAATGCTCTTTAAAATTGAAAA
TGCAACCTGGCAGCGTGTGGTAAGAGCACTTTATCGTAAATACAATAAGGAATTTTTTACATA
TCCAGCCGCCAAAACAAACCACCACGCTTTTGAATCAGGATTGGCATATCACACGGCAACAAT
GGTTCGTTTGGCAGATAGTATCGGAGATATCTATCCAGAACTTAATAAAAGTTTGATGTTTGC
TGGTATTATGCTACATGATTTAGCCAAGGTCATAGAGTTATCGGGTCCTGATAATACAGAATA
TACTATTCGAGGTAATCTTATCGGTCATATTTCACTTATTGATGAGGAATTAA

(B)
LLSNDDKRERYMEQMLFKIENATWQRVVRALYRKYNKEFFTYPAAKTNHHAFESGLAYHTATM
VRLADSIGDIYPELNKSLMFAGIMLHDLAKVIELSGPDNTEYTIRGNLIGHISLIDEEL

ID-52: 436 base pairs
Clone 104
(A)
GTGGTGCCTGTTGAAAATATTTATTTGGATAAACGTATTACGAAGCAAGCTACTCAGTTTTTA
GAGGCTGCTAGAGCAATTGATTCACGAGAACATTTAATTTCGGGTTATCGTAGTGTTGCCTAT
CAGGAGAAGTTGTTCAATTCTTATGTTACTCAAGAGATGACTAGTAACCCTAATTTGACGAGG
GGACAAGCAGAAAAGTTGGTAAAAACTTACTCTCAGCCTGCAGGTGCTAGTGAACACCAGACT
GGATTAGCGATGGATATGAGTACTGTAGATTCTTTGAATGAGAGCGATCCTAGAGTAGTCAGT

CAGTTGAAAAAGATAGCTCCACAATATGGTTTTGTCTTACGGTTTCCGGATGGTAAAACAGCA
GAAACAGGGGTAGGTTATGAAGATTGGCATTACCGCTATGTTGGGGTAGAGTCTGCAAAATAT
ATGGTCAAACATCATTTAA

(B)
MDKRITKQATQFLEAARAIDSREHLISGYRSVAYQEKLFNSYVTQEMTSNPNLTRGQAEKLVK
TYSQPAGASEHQTGLAMDMSTVDSLNESDPRVVSQLKKIAPQYGFVLRFPDGKTAETGVGYED
WHYRYVGVESAKYMVKHHL

ID-53: 190 base pairs
Clone 106
(A)
CTGTTATGTGGATTTCTTCCATCAATTCCTGTGTCTAATTCCGGGGGGTATGGTATAATAACA
GTTATGAAAAATAAAAAAATCTTATTTGGGACTGGCCTTGCTGGTGTGGGTTTACTGGCAGCT
GCTGGTTATACCCTAACTAAAAAAGTAACAGATTATAAACGTCAGCAAATCACTCAGACCTTA
A

(B)
MLCGFLPSIPVSNSGGYGIITVMKNKKILFGTGLAGVGLLAAAGYTLTKKVTDYKRQQITQTL

ID-54: 310 base pairs
Clone 108
(A)
ATGTATCAAACTCAGACAAATAAGGAAAAATTTGTTTTATTTTTGAAATTATTTATCCCAGTA
TTGATTTATCAATTTGCTAATTTTTCAGCTACTTTTATTGATTCGGTTATGACTGGACAGTAT
AGTCAGCTACATTTGGCAGGTGTGTCAACTGCTAGTAATTTATGGACTCCGTTTTTCGCTTTA
TTAGTAGGTATGATTTCAGCATTAGTACCAGTAGTTGGTCAACATTTGGGTAGAGGAAATAAA
GAACAAATTCGCACAGAATTTCATCAATTTCTATATTTAGGTTTGATACTGTCCTTAA

(B)
MYQTQTNKEKFVLFLKLFIPVLIYQFANFSATFIDSVMTGQYSQLHLAGVSTASNLWTPFFAL
LVGMISALVPVVGQHLGRGNKEQIRTEFHQFLYLGLILSL

ID-55: 155 base pairs
Clone 112
(A)
CTGCTCTTTTTAGCTAACTTTTCTAATTTATGGTATAATTGTATGGATTGTTTAGCTAGAATG
GAGAAGATGATGCAAGATGTTTTCATTATAGGAAGTAGAGGGTTGCCAGCTCGTTACGGTGGT
TTTGAAACTTTTGTTTCAGAATTGATTAA

(B)
MLFLANFSNLWYNCMDCLARMEKMMQDVFIIGSRGLPARYGGFETFVSELI

ID-56: 100 base pairs
Clone 120
(A)
TTGAGGAGTAATATGGTAAAGACAGCAGTTTTAATGGCGACATACAATGGCGAAAAATTTATA
TCTGAACAACTTGATTCAATTCGCCAACAGACATTAA

(B)
MRSNMVKTAVLMATYNGEKFISEQLDSIRQQTL

ID-57: 77 base pairs
Clone 123
(A)
GTGATTATGGATAAGTCTATTCCTAAAGCAACTGCTAAACGTTTATCACTGTACTACCGTATT
TTTAAACGTTTTAA

(B)
MIMDKSIPKATAKRLSLYYRIFKRF

ID-58: 476 base pairs
Clone 125
(A)
ATGGGTGCTAAAGGAGCAGATGTCATTCTCGTTTTATCACACTCTGGCATTGGAGATGATCGA
TATGAAGAAGGTGAAGAAAACGTTGGCTATCAAATTGCCAGCATCAAGGGAGTGGATGCCGTT
GTTACGGGACACTCACACGCTGAATTTCCATCAGGTAACGGTACTGGCTTCTATGAAAAATAC
ACTGGAGTTGATGGTATCAATGGAAAAATAAATGGAACACCTGTTACAATGGCAGGCAAGTAC
GGGGATCACCTTGGTATTATTGATTTAGGACTTAGTTATACTAATGGAAAATGGCAAGTCTCC
GAAAGCAGTGCTAAAATCCGTAAAATTGATATGAACTCAACAACTGCTGACGAGCGTATCATT
GCATTGGCTAAGGAAGCACACGATGGCACTATCAACTATGTTCGCCAACAAGTAGGTACAACA
ACTGCGCCAATTACAAGTTACTTTGCACTAGTTAA

(B)
MGAKGADVILVLSHSGIGDDRYEEGEENVGYQIASIKGVDAVVTGHSHAEFPSGNGTGFYEKY
TGVDGINGKINGTPVTMAGKYGDHLGIIDLGLSYTNGKWQVSESSAKIRKIDMNSTTADERII
ALAKEAHDGTINYVRQQVGTTTAPITSYFALV

ID-59: 170 base pairs
Clone 135
(A)
TTGTCAATAAGGTTTCAAATCAGCTTGAAATATGATAAAATAAAACAGATTGTAAGTGACTGT
TTAAGCTTGTTTTTCAGAGAGGTTTTTATGAATACAAACACAATAAAAAAGGTTGTAGCGACT
GGAATTGGAGCTGCACTTTTTATCATTATAGGTATGCTAGTTAA

(B)
MSIRFQISLKYDKIKQIVSDCLSLFFREVFMNTNTIKKVVATGIGAALFIIIGMLV

ID-60: 242 base pairs
Clone 145
(A)
ATGAAACATTTAAAATTTCAATCGGTCTTCGACATTATTGGTCCTGTTATGATTGGACCATCA
AGTAGTCATACTGCAGGAGCTGTCCGCATTGGTAAAGTTGTCCATTCTATTTTTGGTGAACCT
AGTGAAGTAACCTTTCATTTATACAATTCTTTTGCTAAAACTTACCAAGGACACGGTACTGAT
AAAGCATTGGTTGCAGGGATTCTAGGAATGGATACAGATAATCCAGATATTAA

(B)
MKHLKFQSVFDIIGPVMIGPSSSHTAGAVRIGKVVHSIFGEPSEVTFHLYNSFAKTYQGHGTD
KALVAGILGMDTDNPDI

ID-61: 122 base pairs
Clone 147
(A)
GTGTCAGAAGGTGTTTTAATGTTTCTAAAAGAAGATGACGTAGAGACTTTTCTTCATATCCTG
ACAAATTCATTTAGCCAATTTATGGCACAATTTGATTTGTGTCATAAGGAAATGATTAA

(B)
MSEGVLMFLKEDDVETFLHILTNSFSQFMAQFDLCHKEMI

ID-62: 83 base pairs
Clone 150
(A)
ATGACCTACAAAGATTACACAGGTTTAGATCGGACTGAACTTTTGAGTAAAGTGCGTCATATG
ATGTCCGACAAACGTTTTAA

(B)
MTYKDYTGLDRTELLSKVRHMMSDKRF

ID-63: 94 base pairs
Clone S2
(A)
CTGAGTTGGGTCTTGGAAACGGTCCTGTCAATCATACTAGCTATCAAGGAGACTAAAATGTAT
TTAGAACAACTAAAAGAGGTAAATCCTTTAA

(B)
MSWVLETVLSIILAIKETKMYLEQLKEVNPL

This Page Blank (uspto)

FIGURE 3

nucS1

Bgl II Eco RV

5'- cgagatctgatatctcacaaacagataacggcgtaaatag -3'

nucS2

Bgl II Sma I

5'- gaagatcttccccgggatcacaaacagataacggcgtaaatag -3'

nucS3

Bgl II Eco RV

5'- cgagatctgatatccatcacaaacagataacggcgtaaatag -3'

nucR

Bam HI

5'- cgggatccttatggacctgaatcagcgttgtc -3'

NucSeq

5'- ggatgctttgtttcaggtgtatc -3'

$pTREP_F$

5'- catgatatcggtacctcaagctcatatcattgtccggcaatggtgtgggctttttttgttttagcggataa caatttcacac -3'

$pTREP_R$

5'- gcggatcccccgggcttaattaatgtttaaacactagtcgaagatctcgcgaattctcctgtgtgaaatt gttatccgcta -3'

pUC_F

5'- cgccagggttttcccagtcacgac -3'

V_R

5'- tcaggggggcggagcctatg -3'

V_1

5'- tcgtatgttgtgtggaattgtg -3'

V_2

5'- tccggctcgtatgttgtgtggaattg -3'

This Page Blank (uspto)

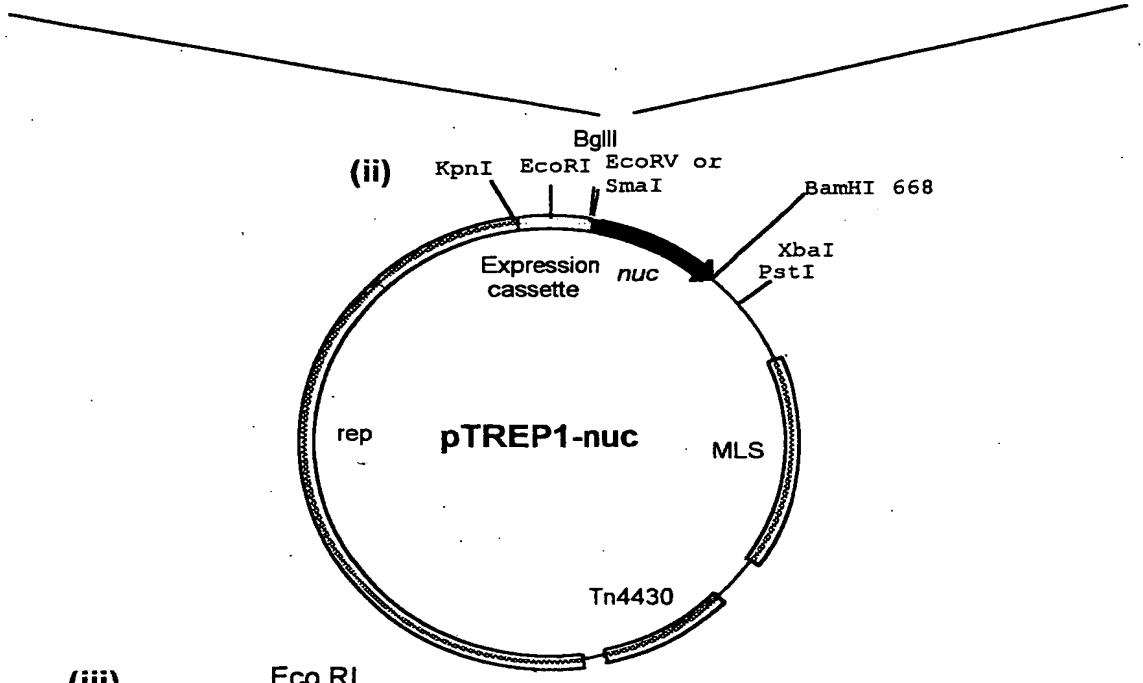
FIGURE 4

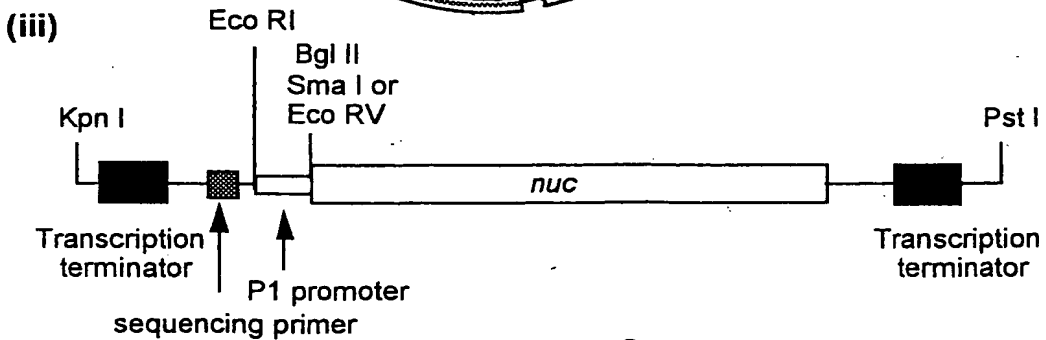
pTREP-Nuc vectors allow cloning of genomic DNA into each frame with respect to the nuclease gene

(i)

```
pTREP1-nuc1 (EcoRV) AAGTATCAGATCT--GATATC--TCACAAACAGATAACGGCGTAAAT Frame =+1
                    :::::::::::::     ▲    :::::::::::::::::::::::::
                    :::::::::::::          :::::::::::::::::::::::::
pTREP1-nuc2 (Sma 1) AAGTATCAGATCTTCCCCGGGA-TCACAAACAGATAACGGCGTAAAT Frame =+2
                    :::::::::::::     ▲    :::::::::::::::::::::::::
                    :::::::::::::          :::::::::::::::::::::::::
pTREP1-nuc3 (EcoRV) AAGTATCAGATCT--GATATCCATCACAAACAGATAACGGCGTAAAT Frame =+3
                                       ▲    :::::::::::::::::::::::::
                                            :::::::::::::::::::::::::
Nuclease Gene                               TCACAAACAGATAACGGCGTAAAT

Cloning site is indicated by an arrow
```





The pTREP-nuc Cassette

09/769,736